



UNIVERSIDADE ESTADUAL DE CAMPINAS
Faculdade de Engenharia Elétrica e de Computação

Isaac Diego Tavares de Souza

Imparcialidade Versus Eficiência na Alocação Espectral em Redes Ópticas Elásticas

Campinas

2017



UNIVERSIDADE ESTADUAL DE CAMPINAS
Faculdade de Engenharia Elétrica e de Computação

Isaac Diego Tavares de Souza

Imparcialidade Versus Eficiência na Alocação Espectral em Redes Ópticas Elásticas

Dissertação apresentada à Faculdade de Engenharia Elétrica e de Computação da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Mestre em Engenharia Elétrica, na Área de Telecomunicações e Telemática.

Orientador: Prof. Dr. Helio Waldman

Este exemplar corresponde à versão final da tese defendida pelo aluno Isaac Diego Tavares de Souza, e orientada pelo Prof. Dr. Helio Waldman

Campinas

2017

Agência(s) de fomento e nº(s) de processo(s): CNPq, 159677/2015-7

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca da Área de Engenharia e Arquitetura
Luciana Pietrosanto Milla - CRB 8/8129

So89i Souza, Isaac Diego Tavares de, 1990-
Imparcialidade versus eficiência na alocação espectral em redes ópticas elásticas / Isaac Diego Tavares de Souza. – Campinas, SP : [s.n.], 2017.

Orientador: Helio Waldman.
Dissertação (mestrado) – Universidade Estadual de Campinas, Faculdade de Engenharia Elétrica e de Computação.

1. Redes ópticas elásticas. 2. Markov, Cadeias de. 3. Imparcialidade. I. Waldman, Helio. II. Universidade Estadual de Campinas. Faculdade de Engenharia Elétrica e de Computação. III. Título.

Informações para Biblioteca Digital

Título em outro idioma: Fairness versus efficiency in spectrum allocation on elastic optical networks

Palavras-chave em inglês:

Elastic optical networks

Markov Chains

Fairness

Área de concentração: Telecomunicações e Telemática

Titulação: Mestre em Engenharia Elétrica

Banca examinadora:

Helio Waldman [Orientador]

Luiz Henrique Bonani do Nascimento

Darli Augusto de Arruda Mello

Data de defesa: 10-08-2017

Programa de Pós-Graduação: Engenharia Elétrica

COMISSÃO JULGADORA - DISSERTAÇÃO DE MESTRADO

Candidato: Isaac Diego Tavares de Souza RA: 162690

Data da Defesa: 10 de Agosto de 2017

Título da Tese: “Imparcialidade Versus Eficiência na Alocação Espectral em Redes Ópticas Elásticas”

Prof. Dr. Helio Waldman (Presidente, FEEC/UNICAMP)

Prof. Dr. Luiz Henrique Bonani do Nascimento (UFABC)

Prof. Dr. Darli Augusto de Arruda Mello (FEEC/UNICAMP)

A ata de defesa, com as respectivas assinaturas dos membros da Comissão Julgadora, encontra-se no processo de vida acadêmica do aluno.

Dedico esta tese a J  ssica.

Agradecimentos

Gostaria de agradecer primeiramente ao professor Helio Waldman que foi imensamente paciente comigo durante todo o meu processo de aprendizado sobre redes ópticas e as habilidades que são requeridas de um pesquisador. Pelo auxílio indispensável no fomento econômico, agradeço ao CNPq. Também sou grato pelos ensinamentos dos demais professores e colegas de laboratório que tive a honra de interagir durante a minha jornada nesta instituição.

Agradeço aos meus pais, à minha tia e aos meus irmãos pelo apoio que veio de diversas formas durante esse período. À minha parceira, Jéssica, nem 1000 páginas de agradecimento seriam suficientes para demonstrar o quanto eu valorizo todo o seu esforço em estar do meu lado durante todos os momentos da construção dessa dissertação.

“We all make choices, but in the end, our choices make us.”
(Ken Levine)

Resumo

O crescimento do tráfego na rede mundial de computadores tem impulsionado a pesquisa em tecnologias que consigam usar a infraestrutura já instalada da forma mais eficiente possível. Essa infraestrutura é, em maior parte, composta de fibra óptica. Para aumentar sua eficiência, foi proposto que a próxima geração de redes ópticas se utilize do paradigma flexível, que aumenta o aproveitamento do espectro em até 40%. Entretanto, a alocação eficiente de recursos gera diversos problemas, entre eles, a discriminação de classes de acesso para tráfego multi-classes. Dessa forma, nesse trabalho, as políticas de alocação espectral que visam a eliminação da discriminação de classes são estudadas e seu desempenho comparado com as políticas que visam somente a eficiência. Assim, o estudo apresenta uma análise quantitativa do *trade-off* entre eficiência espectral e imparcialidade na alocação espectral. Os resultados apontam que o uso de imparcialidade pode gerar ganho no tráfego total agregado médio de um sistema óptico para certas configurações e em carga de tráfego alta.

Palavras-chaves: imparcialidade; redes ópticas elásticas; alocação espectral.

Abstract

The traffic growth on the global computer network have been boosting the research on technologies that are able to use the installed infrastructure in the most efficient way. This infrastructure is, in most part, composed by optical fiber. To improve its efficiency, it was proposed that the next generation of optical networks uses the flexible paradigm, which improves the spectrum use by up to 40%. However, the efficient use of resources generates several problems, between them, the discrimination of access classes on multiclass traffic. Thus, on this work, the spectrum allocation policies that aim to end the discrimination of classes are studied and their performance are compared with the policies that aim only efficiency. Therefore, the study presents a quantitative analysis on the trade-off between spectral efficiency and fairness on spectrum allocation. The obtained results showed that, by applying fairness in the spectrum allocation, the mean total throughput of the network can be increased for certain settings and high traffic load.

Keywords: fairness; elastic optical networks; spectrum allocation.

Lista de ilustrações

Figura 1 – Processo de RMSA.	24
Figura 2 – Alocação espectral para tráfego uniclasse com algoritmo aleatório (a) e First-Fit (b).	27
Figura 3 – Fragmentação do espectro em redes elásticas.	27
Figura 4 – Algoritmo de alocação First-Fit.	28
Figura 5 – Algoritmo de alocação First-Fit com desfragmentação.	30
Figura 6 – Exemplos de particionamento do enlace óptico.	31
Figura 7 – Exemplo de espaço de soluções imparciais do particionamento para duas classes de acesso.	32
Figura 8 – Classificação de estados quanto ao favorecimento de classes.	33
Figura 9 – Exemplos de fragmentação espectral em redes elásticas.	36
Figura 10 – Modelo de enlace único com N slots.	37
Figura 11 – A CTMC para um algoritmo voraz desfragmentado com $\vec{b} = [2, 4]^T$ e $C = \infty$	42
Figura 12 – A CTMC para um algoritmo voraz desfragmentado com $\vec{b} = [2, 4]^T$ e $C = 12$	44
Figura 13 – A CTMC reduzida para o algoritmo <i>First-Fit</i> com $\vec{b} = [2, 4]^T$ e $C = 6$	45
Figura 14 – A CTMC para o algoritmo <i>First-Fit</i> desfragmentado com $\vec{b} = [2, 4]^T$ e $C = 6$	46
Figura 15 – Comparação das CTMCs entre o algoritmo voraz (todos os estados) e o algoritmo de Particionamento (estados dentro do retângulo) com $\vec{b} = [2, 4]^T$, $C = 12$ e $\vec{p} = [4, 8]$	47
Figura 16 – Construção da CTMC do Pseudo-particionamento (com desfragmentação) a partir da CTMC do algoritmo voraz desfragmentado com $\vec{b} = [2, 4]^T$, $C = 12$	48
Figura 17 – Construção da CTMC do <i>Trunk Reservation</i> (com desfragmentação) a partir da CTMC do algoritmo voraz desfragmentado com $\vec{b} = [2, 4]^T$, $C = 12$	49
Figura 18 – CTMC do <i>Trunk Reservation</i> (com desfragmentação) para $\vec{b} = [2, 4]^T$, $C = 12$	50
Figura 19 – Explosão do número de estados numa CTMC de acordo com o número de classes de acesso (todas as classes ocupam um slot e o número de slots totais do sistema fixo em dez).	53
Figura 20 – Fluxograma do funcionamento da ferramenta computacional.	55
Figura 21 – Relação entre as amostras de vazão média e de PB.	60
Figura 22 – Interpolação dos valores de Vazão para os valores de PB não presentes nos resultados originais.	60
Figura 23 – Vazão por carga de tráfego para $S = 164$ slots.	63

Figura 24 – Comparação do custo da imparcialidade em vazão dos dois algoritmos imparciais para $S = 164$ slots.	64
Figura 25 – Probabilidade de bloqueio dos algoritmos de alocação analisados para $S = 164$ slots.	65
Figura 26 – Vazão por PB para $S = 164$ slots.	66
Figura 27 – Comparação do custo da imparcialidade para vazão por PB $S = 164$ slots. . .	66
Figura 28 – Ocupação por carga de tráfego para $S = 164$ slots.	67
Figura 29 – Imparcialidade por carga de tráfego para $S = 164$ slots.	69
Figura 30 – Vazão por carga de tráfego para $S = 163$ slots.	70
Figura 31 – Comparação do particionamento para os dois tamanhos de espectro analisados.	70
Figura 32 – Comparação do custo da imparcialidade em vazão para $S = 163$ slots. . . .	71
Figura 33 – Probabilidade de bloqueio dos algoritmos de alocação analisados para $S = 163$ slots.	71
Figura 34 – Comparação do custo da imparcialidade para vazão por PB $S = 163$ slots. . .	72
Figura 35 – Comparação do custo da imparcialidade para vazão por PB $S = 163$ slots. . .	73
Figura 36 – Ocupação por carga de tráfego para $S = 163$ slots.	73
Figura 37 – Imparcialidade por carga de tráfego para $S = 163$ slots.	74
Figura 38 – Vazão versus carga de tráfego para $S = 163$ slots (Pseudo-particionamento)	75
Figura 39 – Comparação das partições entre o Particionamento e o Pseudo-particionamento para $S = 163$ slots.	75
Figura 40 – Índice de imparcialidade de Jain para $S = 163$ slots (Pseudo-particionamento)	76
Figura 41 – Cadeia de Markov para altas carga de tráfego utilizando um algoritmo voraz de alocação para $\vec{b} = [2, 4]^T$, $C = 12$ slots.	78
Figura 42 – Cadeia de Markov para altas carga de tráfego utilizando um algoritmo de particionamento para $\vec{b} = [2, 4]^T$, $C = 12$ slots.	79
Figura 43 – Cadeia de Markov para altas carga de tráfego utilizando um algoritmo TR para $\vec{b} = [2, 4]^T$, $C = 12$ slots.	79
Figura 44 – Cadeia de Markov para altas carga de tráfego utilizando um algoritmo voraz de alocação para $\vec{b} = [13, 28]^T$, $C = 164$ slots e $C = 163$ slots. Os estados e transições em tracejados são exclusivos do caso em que $C = 164$ slots. . .	81
Figura 45 – Cadeia de Markov para altas carga de tráfego utilizando um algoritmo de particionamento para $\vec{b} = [13, 28]^T$, $C = 164$ slots e $C = 163$ slots. Os estados e transições em tracejados são exclusivos do caso em que $C = 164$ slots.	82
Figura 46 – Cadeia de Markov para altas carga de tráfego utilizando um algoritmo TR para $\vec{b} = [13, 28]^T$, $C = 164$ slots.	83
Figura 47 – Custo da imparcialidade assintótico para $\vec{b} = [13, 28]^T$ e $S = 164$ slots. . .	84
Figura 48 – Custo da imparcialidade assintótico para $\vec{b} = [13, 28]^T$ e $S = 163$ slots. . .	84

Lista de tabelas

Tabela 1 – Valores mínimos dos parâmetros de um transceiver óptico elástico para alcance transparente de 8000 km.	25
Tabela 2 – Configurações dos conjuntos de resultados da Seção 4.3	62

Lista de Acrônimos e Abreviações

BVT	Bandwidth Variable Transceiver
CMTC	Continuous Time Markov Chain
Co-OFDM	Coherent Optical Frequency Division Multiplexing
CoWDM	Coherent Wavelength Division Multiplexing
DP	Dedicated Partitioning
ITU	International Telecommunication Union
ITU-T	Telecommunication Standardization Sector on ITU
MLR	Mixed-Line-Rate
OAWG	Optical Arbitrary Waveform Generation
PB	Probabilidade de Bloqueio
QAM	Quadrature and Amplitude Modulation
RMSA	Routing, Modulation and Spectrum Assignment
RSA	Routing and Spectrum Assignment
RWA	Routing and Wavelength Assignment
TR	Trunk Reservation
TRR	Two Rate Reservation
WDM	Wavelength Division Multiplexing
WSS	Wavelength Selective Switch

Sumário

1	Introdução	16
1.1	Exposição do problema	16
1.2	Organização do trabalho	17
1.3	Publicações relacionadas	18
I	Fundamentos	22
2	Algoritmos de RMSA	23
2.1	Routing, Modulation and Spectrum Assignment	23
2.1.1	Roteamento	24
2.1.2	Modulação	25
2.1.3	Alocação espectral	26
2.2	Desfragmentação	28
2.2.1	Técnicas de desfragmentação	28
2.2.2	Alocação espectral e desfragmentação	29
2.3	Imparcialidade e o seu custo	30
2.3.1	Particionamento	31
2.3.2	Trunk Reservation	32
2.3.3	Custo da imparcialidade	33
2.4	Métricas de Avaliação de Desempenho	34
2.4.1	Tráfego médio alcançado	34
2.4.2	Ocupação média	34
2.4.3	Probabilidade de bloqueio	34
2.4.4	Índice de Imparcialidade	35
2.5	Fragmentação e enlace único	35
II	Contribuições	38
3	Modelo analítico	39
3.1	Modelagem do sistema	39
3.2	Desfragmentação e Reversibilidade	44
3.3	Algoritmos baseados em políticas não-vorazes	46
3.3.1	Particionamento	47
3.3.2	Trunk Reservation	49
3.4	Limitações do método analítico	52
4	Custo da Imparcialidade	54
4.1	Ferramenta computacional	54

4.1.1	Funcionamento da ferramenta	54
4.1.2	Cálculo das métricas de desempenho	56
4.1.3	Obtenção dos Resultados Gráficos	58
4.2	Arranjo Experimental	61
4.3	Resultados e análises	62
4.3.1	Conjunto de resultados 1 (164 slots)	62
4.3.2	Conjunto de resultados 2 (163 slots)	69
4.3.3	Conjunto de resultados 3 (Pseudo-particionamento)	74
5	Comportamento assintótico	77
5.1	Efeito de aumento do tráfego nas cadeias de Markov	77
5.2	Maximização da vazão	80
	Conclusão	85
	 Referências	 87

1 Introdução

1.1 Exposição do problema

A rede mundial de computadores tem apresentado um aumento em número de usuários cada vez maior no decorrer dos últimos anos. Esse aumento se deve à introdução de novas tecnologias que permitiram não só que mais usuários humanos entrassem na rede, mas também que os próprios computadores se comunicassem cada vez mais entre si.

Para acompanhar esse aumento, a infra-estrutura da rede tem sido expandida com o tempo para agregar o tráfego adicional. Essa infra-estrutura tem como base as fibras ópticas. Esse tipo de material foi escolhido pelas suas excelentes propriedades físicas, sendo considerado o material mais adequado para tráfego de grandes quantidades de informação a longas distâncias.

A inserção de novas unidades de fibra óptica na rede é uma operação de custos elevados que tende a ser evitada ao máximo pelos agentes provedores dessa infra-estrutura. Dessa forma, existe uma constante necessidade de se aumentar o aproveitamento da fibra já instalada, motivando uma pesquisa de caráter multidisciplinar que envolve diversas técnicas de engenharia de telecomunicações como modulação e roteamento, mas também de economia, marketing e etc.

A tecnologia de comunicação por fibras ópticas, motivada pelos desafios citados anteriormente, passou, então, por diversas melhorias. Como descrito em (AGRAWAL, 2012), a evolução das comunicações ópticas pode ser dividida em três eras. A primeira era foi marcada pela detecção direta e o uso de regeneradores. A segunda era foi marcada por sistemas de *long-haul* não regenerados e sistemas WDM (*Wavelength Division Multiplexing*). A terceira era, na qual nos situamos nesse momento, começou com a introdução da tecnologia de detecção coerente. Um desafio dessa última é utilizar o espectro óptico de forma mais eficiente e flexível, ou seja, agregar o tráfego da maior quantidade de usuários possível, dando liberdade para suas demandas serem distintas, considerando gerar o mínimo de perda de aproveitamento do espectro.

Para se alcançar esse objetivo, o paradigma das redes WDM, baseado em uma grade fixa de slots espectrais, deve ser substituído pelo paradigma flexível (ou elástico), no qual a grade de slots não mais obedece a uma tabela fixa, mas sim permite conexões de qualquer tamanho serem alocadas em qualquer região do espectro não ocupado.

Essa tecnologia vem como resposta a diversas demandas geradas do aumento do volume e heterogeneidade do tráfego de dados. A difusão de linhas com múltiplas taxas (ou MLR, *Mixed-Line-Rate*) originada do uso elevado de *grooming* (NAG *et al.*, 2010) em redes ópticas requer mais do que as redes WDM legadas podem oferecer. O novo paradigma de computação que está migrando a capacidade de processamento para os *data centers*, demanda o uso de redes

elásticas, como é demonstrado em (KLINKOWSKI; WALKOWIAK, 2013).

O problema de alocação de *lightpaths* no espectro, nas redes WDM, se resume à escolha de slots que seja suficiente para acomodar a demanda e que estejam disponíveis nas mesmas posições do espectro em cada enlace do *lightpath*. Os algoritmos que resolvem esse problema são chamados de RWA (*Routing and Wavelength Assignment*). Entretanto, nas redes ópticas elásticas, o problema de alocação se torna RSA (*Routing and Spectrum Assignment*), pois, requisições não serão designadas a canais padronizados seguindo uma grade fixa, mas sim, os canais serão especificados conforme a banda demandada, podendo tomar posição em qualquer região do espectro óptico em que esta banda estiver disponível.

Os gestores desse *backbone* óptico podem apresentar diversas estratégias quanto à comercialização do seu recurso, como, por exemplo, agregar a maior quantidade de clientes que a sua infraestrutura permitir ou maximizar a vazão de dados sendo transmitida na rede. Geralmente, esses objetivos são alcançados utilizando algoritmos de RSA que usem o espectro de forma mais eficiente.

A maximização do uso do recurso oferecido é só uma das possíveis estratégias de comercialização desse produto. A distribuição imparcial do espectro entre as diferentes categorias de clientes, garantindo que o clientes de maior tráfego não vão experimentar uma maior dificuldade em acessar a infraestrutura, pode ser um diferencial do serviço ou até um requisito de alguns clientes.

Dessa forma, o trabalho descrito a seguir tem como principais objetivos analisar diversos algoritmos de alocação espectral, avaliando o compromisso entre maximização da vazão de dados na rede e a imparcialidade da alocação entre as diferentes classes de acesso.

1.2 Organização do trabalho

Após a breve introdução ao problema presente nesse capítulo, o restante desse trabalho é composto de duas partes: a fundamentação teórica e as contribuições da pesquisa.

A fundamentação teórica inicia no Capítulo 2, com os conceitos de roteamento, modulação e alocação espectral, as implicações do novo paradigma de redes ópticas nas técnicas de alocação já conhecidas, o conceito de imparcialidade na alocação de conexões no espectro óptico e como avaliar o desempenho dessas técnicas. Já no Capítulo 3, é mostrada a fundamentação matemática na qual os algoritmos de alocação foram simulados e avaliados.

Na segunda parte do trabalho, a primeira contribuição é a comparação entre os algoritmos de alocação clássicos e os algoritmos imparciais, presente no Capítulo 4. A outra contribuição dessa pesquisa, que pode ser vista no Capítulo 5, é a análise do comportamento do algoritmo de *Trunk Reservation* quando a carga de tráfego geral tende a valores altos.

Por fim, na Capítulo de Conclusões, é feita uma síntese de todas as conclusões obtidas

dos resultados e são avaliadas as perspectivas de trabalhos futuros que essas conclusões abrem.

1.3 Publicações relacionadas

Uma visão geral das motivações e desafios tecnológicos envolvendo o advento das redes ópticas elásticas é mostrada em (GERSTEL *et al.*, 2012). Nesse artigo, é mostrado que o paradigma de redes ópticas elásticas é impulsionado por requisitos não cumpridos pelo paradigma atual de grid fixo. Esses requisitos são: suportar taxas de transmissão superiores a 100 Gb/s; suportar diversidade de largura de banda entre as diferentes conexões; diminuir os espaçamentos entre os canais; ajustar o formato de modulação para cumprir os requisitos da demanda eficientemente; permitir que a camada óptica se comunique diretamente com a camada do cliente. Ele indica que a implantação do paradigma pode levar a ganhos de até 150% no aproveitamento do espectro. Também é mostrado que os principais componentes possibilitadores do paradigma são transmissores/receptores de largura de banda flexível (ou BVT, *Bandwidth Variable Transceiver*) e switches seletivos de espectro (ou flex WSS, *Wavelength Selective Switch*). O artigo ainda mostra que estão em análise diversas soluções multiportadora para implementação dos transceivers da próxima geração. São elas: CoWDM (*Coherent Wavelength Division Multiplexing*), Co-OFDM (*Coherent Optical Frequency Division Multiplexing*), Nyquist-WDM (*Wavelength Division Multiplexing*) e OAWG (*Optical Arbitrary Waveform Generation*) dinâmico.

Em (TOMKOS *et al.*, 2014), é explicado que, com a chegada das redes ópticas elásticas, as técnicas clássicas de roteamento e alocação de comprimento de onda (RWA) não mais podem ser aplicadas, pois restrições como as posições fixas no espectro e ocupação predeterminada de cada conexão deixam de existir. Nas redes elásticas, outros problemas surgem, como por exemplo, o fato de que conexões de mais de um slot (unidade mínima na divisão do espectro) terem que ser alocadas em um conjunto contíguo de slots. Por esses motivos, as técnicas aplicadas a esse tipo de rede são chamadas de roteamento e alocação de espectro (ou de RSA).

O objetivo desse trabalho é a caracterização do compromisso entre eficiência e imparcialidade em redes ópticas elásticas. O “preço” da imparcialidade é discutido de forma abrangente em (BERTSIMAS *et al.*, 2011). O trabalho justifica esse estudo no fato de que as soluções eficientes não serem sempre implementadas na prática. Isso acontece por causa do desequilíbrio social gerado pelo favorecimento de uma porção dos usuários do recurso compartilhado. É realizada, então, a análise desse “preço” utilizando e comparando duas definições distintas de imparcialidade: *max-min* (BONALD; MASSOULIÉ, 2001) e proporcional (KELLY *et al.*, 1998). O trabalho conclui que o conceito de imparcialidade proporcional gera menor custo, principalmente quando número de usuários distintos é grande. Por isso, a definição de imparcialidade utilizada nesse trabalho tem semelhanças com a imparcialidade proporcional.

Ainda no tópico de imparcialidade versus eficiência, em (ZUKERMAN *et al.*, 2005) é introduzido o conceito de imparcialidade α , que é uma medida da imparcialidade de um al-

goritmo de alocação. Assim, utilizando essa medida como “entrada”, os autores criam curvas relacionando a arrecadação com a venda do espectro com o nível de imparcialidade desejado. Ele mostra que o modelo matemático utilizado pode ser aplicado para diversas topologias, critérios de imparcialidade e modelo de negócio. Nessa dissertação, uma análise semelhante foi realizada relacionando o grau de imparcialidade com a carga de tráfego na rede, para um algoritmo semi-imparcial.

Enfatizando a análise do desempenho de algoritmos de alocação espectral no enlace único, em (WALDMAN *et al.*, 2015) tem-se a identificação do algoritmo de desfragmentação como o limite superior em termos do aproveitamento do espectro. São analisados 3 algoritmos de alocação espectral em comparação com primeiro: Aleatório, *First-Fit* e *Bilateral First-Fit*. O algoritmo Aleatório apresenta maiores perdas por fragmentação, enquanto os outros dois apresentam resultados semelhantes. Nesse trabalho são definidos como algoritmos vorazes aqueles que necessariamente alocam o recurso demandado, caso esteja disponível. Com o uso da desfragmentação, todos os algoritmos vorazes apresentam o mesmo desempenho.

Na análise dos métodos imparciais de alocação espectral, em (CALLEGATI *et al.*, 2013) é feito um estudo da imparcialidade usando como critério a diferença nas probabilidades de bloqueio entre as diferentes classes de acesso e o impacto da fragmentação do espectro nesse quesito. Continuando este trabalho, em (CALLEGATI *et al.*, 2014) é explicada a política de alocação de *Trunk Reservation* (ou TR). Esta política pode ser aplicada a um algoritmo de alocação parcial de forma a torná-lo imparcial. Ela consiste em bloquear conexões que requerem menos slots que outras, em situações em que ocorreria discriminação das classes, de forma a não prejudicar as que requerem mais slots. O algoritmo de TR é um dos algoritmos imparciais estudados nessa dissertação.

Ainda nos métodos imparciais de alocação, (WANG; MUKHERJEE, 2014) traz a análise de dois algoritmos de particionamento em comparação com o algoritmo de compartilhamento total do espectro. O primeiro algoritmo proposto é chamado Pseudo-Particionamento, no qual as classes de acesso são divididas em dois grupos e as conexões de cada grupo são alocadas nas duas “fronteiras” do espectro: as conexões que requerem menos slots, no início do espectro e as conexões que requerem mais slots no final. O segundo algoritmo proposto é chamado Particionamento Dedicado ou DP (*Dedicated Partitioning*). Nele, cada classe tem uma partição própria que deve ser determinada de forma otimizada de acordo com o padrão de tráfego. Os resultados mostram que o algoritmo DP é melhor que os outros dois em termos de menor geração de perdas por fragmentação e de melhor imparcialidade sob tráfego uniforme. O desempenho de ambos os algoritmos de particionamento foram analisados nessa dissertação.

Em (WANG *et al.*, 2014), um algoritmo de alocação semi-imparcial é proposto, o TRR (*Two Rate Reservation*). Ele tem uso no caso de duas classes de acesso com tamanho de conexão diferentes. O algoritmo determina que devem existir 3 partições, duas de uso exclusivo de cada classe e uma extra para as duas classes utilizarem conjuntamente. É mostrado que algo-

ritmo aumenta a imparcialidade, baseando-a na razão entre as probabilidades de bloqueio. Os resultados são expandidos em (KIM *et al.*, 2015), onde é mostrado que, mesmo para o caso com taxas de chegadas diferentes entre as classes de acesso, o algoritmo ainda gera resultados mais imparciais que os algoritmos convencionais. O TRR pode ser visto como um caso particular do particionamento em que existem somente duas classe de acesso.

O desempenho dos algoritmos de alocação espectral mais comuns é estudado em (CHATTERJEE; OKI, 2015). Ele é expresso em termos de complexidade, probabilidade de bloqueio e razão de slots contíguos alinhados disponíveis. Os algoritmos analisados são: *First fit*, *Random fit*, *First-last fit*, *Least Used*, *Most Used* e *Exact fit*. O algoritmo *First-last fit* possui desempenho superior a todos os outros analisados em todos os quesitos estudados. O artigo não faz uso da técnica de desfragmentação.

Ainda no estudo dos algoritmos de alocação espectral, em (CHATTERJEE *et al.*, 2016) é proposto um algoritmo de alocação eficiente com o intuito de reduzir os efeitos da fragmentação do espectro em redes elásticas chamado *First-Last-Exact-Fit*. Para isso, o algoritmo divide as conexões em dois grupos: um grupo para as que podem ser alocadas usando apenas um enlace e um outro grupo para as conexões que só podem ser alocadas com dois ou mais enlaces. Dessa forma, ele aloca as conexões do primeiro grupo usando o algoritmo *First-Exact-Fit* e as conexões do segundo grupo usando o algoritmo *Last-Exact-Fit*. Os resultados mostram que, em comparação com os algoritmos tradicionais como *First-Fit* e *First-Last-Fit*, a fragmentação geral do espectro é reduzida e, conseqüentemente, o aproveitamento do espectro é maior. Essa lógica foi considerada nessa dissertação como generalização do algoritmo de particionamento para duas partições e mais de três classes de acesso.

O primeiro modelo analítico de ocupação da rede óptica elástica foi proposto em (YU *et al.*, 2014). Nele, é feita uma análise preliminar de todos requerimentos das redes elásticas e são explicitadas suas diferenças para com o modelo de redes WDM. O autor sugere algumas métricas de desempenho dos algoritmos de alocação espectral, com destaque para a definição da taxa de fragmentação como o proporcional ao número de pares de slots contíguos que estão em estados diferentes (ocupado/desocupado). Em seguida, o autor mostra em detalhes seu modelo e sugere algoritmos heurísticos para o cálculo eficiente das probabilidades de estado. Os resultados obtidos se aproximam dos simulados com pequena margem de erro e essa diferença é atribuída ao uso de heurísticas no cálculo. Para manter os resultados estatisticamente precisos, nessa dissertação não foram utilizadas heurísticas, entretanto é reconhecido o impacto do cálculo direto na complexidade computacional da solução apresentada.

Outro modelo analítico da ocupação do espectro foi proposto em (KIM *et al.*, 2016). Nesse artigo, o espectro é dividido em um grid “semi-flexível”. Nessa abordagem, as conexões não podem ser alocadas em todos as possíveis posições do espectro, mas somente em algumas pré-selecionadas, o que reduz ou até zera o impacto da fragmentação. Em seguida, os autores utilizam o algoritmo de alocação aleatória para comparar os resultados analíticos obtidos com

o modelo e os resultados obtidos por simulação de eventos. As comparações mostram pouca divergência nos resultados. A metodologia analítica, aplicada nesse e no artigo anterior, foi também utilizada nessa dissertação.

Parte I

Fundamentos

2 Algoritmos de RMSA

Esse capítulo trata da contextualização teórica sobre os tópicos relacionados a Roteamento, Modulação e Alocação de Espectro, do inglês *Routing, Modulation and Spectrum Assignment* ou somente RMSA. Essa contextualização se faz necessária para o entendimento dos algoritmos de alocação utilizados no Capítulo 4.

Assim, esse capítulo se configura da seguinte forma: na Seção 2.1 é explicado o conceito de RMSA, as principais técnicas e os desafios encontrados no desenvolvimento dessas técnicas; em seguida, na Seção 2.2, é mostrada a técnica de desfragmentação do espectro e como o seu uso altera o desempenho dos algoritmos de alocação espectral; já na Seção 2.4, são mostradas as principais técnicas de avaliação de desempenho desses algoritmos; por fim, na Seção 2.5, são mostradas as vantagens do estudo do enlace único, preliminarmente ao estudo de uma rede mais complexa.

2.1 Routing, Modulation and Spectrum Assignment

No contexto de operação de uma rede óptica, diversos protocolos devem ser seguidos pelos agentes da rede de forma que o fluxo de informação seja o mais otimizado possível, reduzindo ao máximo os problemas de acesso concorrente ao meio de transmissão. A definição dessa série de protocolos deve atender diversos requisitos e abrange diversas camadas de operação da rede, desde o nível mais baixo, na escolha dos parâmetros de transmissão da luz no meio físico (fibra) até as camadas mais superiores de controle de acesso ao meio, de tráfego da rede e de aplicação.

Para as redes ópticas WDM, essa definição é denominada o problema de Roteamento e Alocação de Comprimento de Onda, ou RWA. O roteamento é a operação de escolha da melhor rota na rede para a transferência de dados acontecer. Essa operação é explicada em mais detalhes posteriormente na Seção 2.1.1. Já a operação de Alocação de Comprimento de Onda é a escolha do melhor comprimento de onda que deverá ser utilizado para transmissão dos dados, de forma que esse não deve ser igual ao de outra conexão já estabelecida em qualquer um dos enlaces do caminho selecionado pela operação anterior. Essa escolha também deve facilitar a alocação de novas conexões.

Com o advento das redes ópticas elásticas, os algoritmos que solucionam o problema RWA não se aplicam com o mesmo desempenho do observado anteriormente. Isso acontece porque, no paradigma elástico, as conexões não tem frequências e tamanhos fixos e pré-definidos para serem posicionadas no espectro. Assim, um novo tipo de problema surge, o de Roteamento e Alocação de Espectro, ou RSA. O roteamento e alocação de espectro são explicados em mais

detalhes nas seções 2.1.1 e 2.1.3.

Por fim, uma das características das redes ópticas elásticas que indica que o desempenho desse tipo de rede é superior ao paradigma anterior é a escolha de como modular o sinal de acordo com a distância que a conexão deseja alcançar. Isso permite um uso mais inteligente dos equipamentos, reduzindo a ocupação espectral sempre que possível mediante o uso eficiente da energia de acordo com o alcance demandado pela rede. A escolha do formato de modulação é explicada em mais detalhes na Seção 2.1.2.

A Figura 1 mostra uma descrição visual da ligação entre as operações que serão discutidas nessa seção.

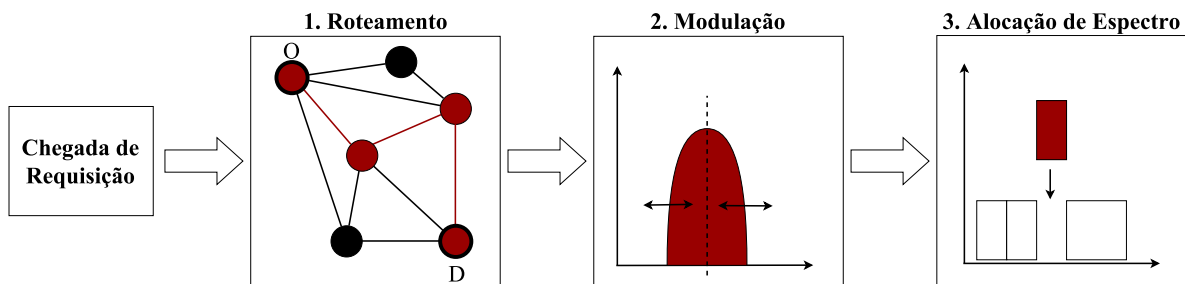


Figura 1 – Processo de RMSA.

2.1.1 Roteamento

Embora não seja o escopo dos resultados mostrados nesse trabalho, a definição do problema de roteamento é bastante significativa no contexto de gerenciamento de redes e pode interferir substancialmente nos resultados.

O problema de roteamento pode ser visto como um problema do menor caminho em um grafo. Os nós do grafo representam as interfaces de conexão da rede, nas quais o tráfego pode ser multiplexado e demultiplexado para que seja distribuído em redes de menor escala. As arestas do grafo podem ser entendidas como os enlaces da rede. Diversas técnicas para a solução desse problema já foram exploradas na literatura. Como mostrado em (BERTHOLD *et al.*, 2008), uma técnica bastante comum para a solução desse problema é o algoritmo de Dijkstra (DIJKSTRA, 1959).

Como explicado em (RUAN; ZHENG, 2014), existe ainda a possibilidade de utilizar a técnica de *multi-path routing*. Nesse tipo de abordagem, para um requisição de banda B , o algoritmo de roteamento deve encontrar $N \geq 2$ caminhos disjuntos em enlaces que tenham banda agregada de, pelo menos, B . Assim, o algoritmo distribui a banda igualmente entre os caminhos encontrados.

2.1.2 Modulação

Como proposto por (GERSTEL *et al.*, 2012), as redes ópticas elásticas podem se utilizar da propriedade de oferecer ocupação não predefinida para gerar ocupações adaptáveis para cada situação que as requisições de conexão apresentam. Por exemplo, a dois clientes da rede que requisitam conexões com a mesma taxa de bit podem ser alocados quantidades de slots diferentes. Isso pode acontecer porque a distância de cada conexão pode diferir, levando a formatos de modulação diferentes. Em geral, formatos de modulação que possuem maior ocupação espectral são menos susceptíveis a perdas com a distância do que formatos de menor ocupação (mais eficientes).

O componente da rede óptica que trata da transformação do sinal digital (dados) do domínio elétrico para o domínio óptico é chamado de transceptor. Os transceptores ópticos elásticos possuem a característica de ter diversos parâmetros que podem ser selecionados para a definição da onda que se propagará pelo meio físico.

Um desses parâmetros é o formato de modulação do sinal. O formato de modulação pode apresentar níveis crescentes de bits por símbolo, cuja a escolha impacta o alcance do sinal sem repetidores, também chamado de alcance transparente. Em geral, quanto maior a quantidade de bits por símbolo do formato de modulação menor será o alcance transparente do sinal. Essa diminuição no alcance ocorre porque a agregação de bits por símbolo torna o sinal mais susceptível à ação da degradação do sinal quando transmitido na fibra. O envio e a recepção de sinais mais complexos também demanda um consumo maior de energia nos transceptores, mantido o mesmo alcance transparente.

Assim, ao receber a requisição de uma conexão na rede, deseja-se que o sistema seja capaz de fazer a escolha do formato de modulação que alcance a distância requerida minimizando os custos envolvidos no processo, por exemplo, a ocupação espectral ou o consumo de energia, se necessário.

Esse compromisso é estudado em (SOUZA *et al.*, 2016), em que é feita uma análise do impacto da seleção de diversos parâmetros no desempenho de transceptores ópticos elásticos. A tabela 1 mostra, para uma requisição de 8000 km, os parâmetros mínimos que dever ser utilizados para alcançar uma transmissão com taxa de erro de bit da ordem de 10^{-15} .

Tabela 1 – Valores mínimos dos parâmetros de um transceiver óptico elástico para alcance transparente de 8000 km.

Taxa de bit (GB/s)	FM	R_s (GBd)	OH (%)	#Slots	#Portadoras
400	8-QAM	32	92	13	4
1000	8-QAM	32	73	28	9

Na tabela, FM é a sigla para o formato de modulação; QAM é o formato de modulação *Quadrature and Amplitude Modulation* e o número que o precede descreve a ordem do formato de modulação; R_s é a taxa de símbolos; OH é o *overhead* adicionado pelo código de correção

de erro em porcentagem do tamanho do *payload* original; #Slots é número de slots de 12,5 GHz que a classe de acesso em questão ocupa por conexão; e #Portadoras é número de portadoras necessárias para o funcionamento de uma conexão da classe de acesso.

2.1.3 Alocação espectral

A alocação de espectro é o conjunto de técnicas que definem como as novas requisições serão alocadas no espectro. Em redes ópticas convencionais (WDM), a banda de operação do espectro óptico geralmente é a banda C (1530 a 1565 nm de comprimento de onda). O espectro da banda de interesse foi, então, dividido em blocos de tamanho igual, para serem utilizados na alocação de conexões. Os blocos tem tamanho predefinido dependendo da padronização seguida. A *International Telecommunication Union* (ITU), em seu braço de padronização de telecomunicações conhecido como *Telecommunication Standardization Sector* (ou ITU-T), padronizou as redes WDM com canais de 50 GHz.

Dessa forma, as requisições de conexão devem ser atendidas pelo algoritmo de alocação espectral que irá posicionar as conexões no espectro dentro dos blocos não ocupados por outras conexões. Esses blocos alocados devem ser os mesmos em todas os enlaces da rede. Essa condição é chamada restrição de continuidade.

Com a chegada do modelo de rede óptica flexível, o ITU-T padronizou a divisão do espectro em blocos (ou *slots*, como serão chamados adiante) de 12,5 GHz. Diferentemente das redes WDM, ou de grid fixo, as conexões nas redes elásticas podem ocupar mais de um slot, contanto que sejam consecutivos no espectro (restrição de contiguidade).

Durante o funcionamento de uma rede não ocorrem somente chegadas de novas conexões, mas também, o término das conexões atualmente alocadas. Essa saída de conexões deixa slots não alocados, que podem ser reaproveitados por novos clientes da rede. Esse processo de geração de “buracos” no espectro é chamado de fragmentação espectral. Para a alocação de usuários com requisitos de tráfego não uniforme, haverá formação de vacâncias que não podem ser ocupadas por nenhuma classe de acesso (Figura 2.a), acarretando perda por fragmentação. Entretanto, caso todos os clientes solicitem conexões iguais em termos de ocupação espectral (tráfego uniclasse) e o algoritmo de alocação as insere em posições previamente escolhidas (Figura 2.b), a fragmentação não acarreta perda de desempenho, dado que uma nova conexão sempre pode ocupar o espaço deixado por uma outra. Entretanto, a uniformidade entre as classes de acesso (tipos de cliente) pode não ocorrer quando a rede em questão utiliza-se do paradigma elástico. Nesse caso, quando os clientes da rede não solicitam quantidades uniformes de slots por conexão (tráfego multiclasse), a fragmentação pode levar ao bloqueio de novas conexões, gerando perda de desempenho (perda de fragmentação) (WALDMAN *et al.*, 2015). A Figura 3 descreve uma situação de fragmentação advinda da diversidade nas classes de acesso.

O sistema se encontra inicialmente no estado (a), com uma conexão que ocupa dois

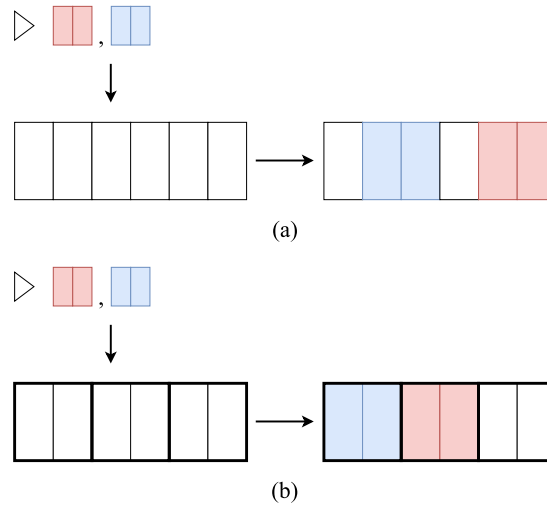


Figura 2 – Alocação espectral para tráfego uniclasse com algoritmo aleatório (a) e First-Fit (b).

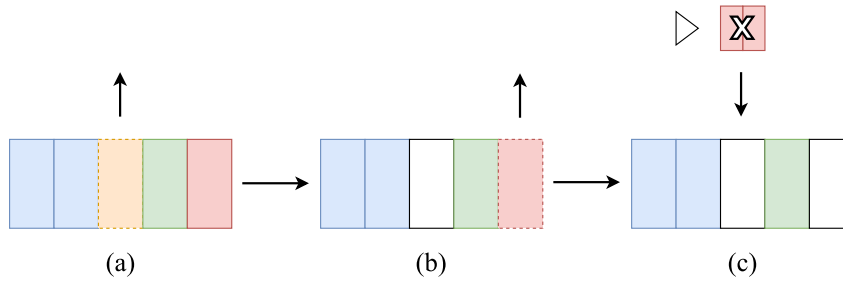


Figura 3 – Fragmentação do espectro em redes elásticas.

slots e três conexões que ocupam um slot cada uma. Um conexão de um slot termina, levando o sistema ao estado (b). Depois disso, outra conexão de um slot termina, levando os sistema ao estado (c). Se, nesse momento, ocorre a chegada de uma requisição por uma conexão de 2 slots, essa deverá ser recusada pois não há slots contíguos disponíveis para a alocação da mesma.

Dessa forma, também é função do algoritmo de alocação evitar situações como a descrita na Figura 3. Isso pode ser feito adaptando-se algoritmos de alocação para diminuir as chances de o sistema chegar a estados bloqueantes por fragmentação. Uma segunda abordagem, mais robusta porém computacionalmente complexa, é a desfragmentação do espectro.

Diversas técnicas de alocação espectral para as redes elásticas foram propostas na literatura, como foi sumarizado no artigo tutorial de (TOMKOS *et al.*, 2014). Em (WALDMAN *et al.*, 2015), foi mostrado que, para o enlace único, qualquer algoritmo de alocação voraz (maximização da ocupação imediatamente após a alocação) apresenta o mesmo desempenho quando utilizada a desfragmentação do espectro.

Dentre os diversos algoritmos vorazes, um exemplo simples é o algoritmo de *First-Fit*. Esse algoritmo tem como lógica de alocação o posicionamento das conexões o mais próximo possível do “início” (ou fim) do espectro, ou seja, nos comprimentos de onda de menor (ou maior) valor, mantendo as conexões agrupadas numa região só. Assim, qualquer nova cone-

xão será alocada no primeiro conjunto de slots contíguos vazios que tiver quantidade de slots suficiente para acomodar uma conexão da classe solicitante.

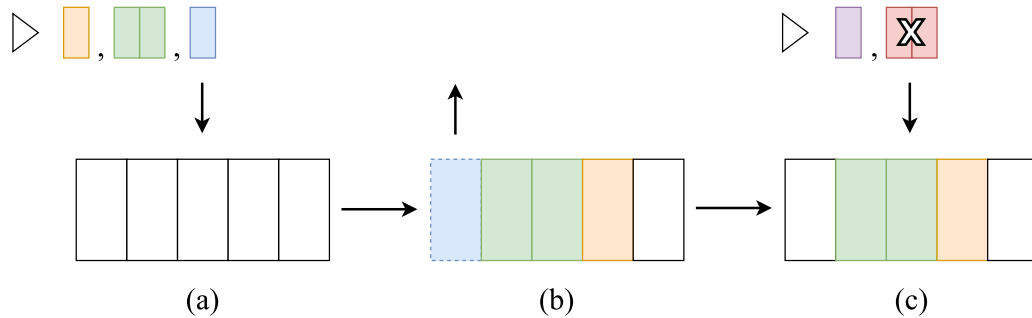


Figura 4 – Algoritmo de alocação First-Fit.

A Figura 4 traz um exemplo de funcionamento desse algoritmo. O sistema exemplificado aqui inicia sua operação sem nenhuma conexão alocada (a). Nesse estado, ele recebe requisições de conexões de duas classes de acesso, uma que requisita um slot e outra que requisita dois. O algoritmo às aloca por ordem de chegada nos slots de menor índice o que leva o sistema ao segundo estado (b). Nesse estado, a primeira conexão alocada termina, tornando vago um slot no início do espectro e levando ao último estado da Figura (c). Nesse estado, se uma conexão de dois slots chegar, será bloqueada pelo algoritmo, mesmo havendo slots suficientes na rede. Isso acontece porque esses dois slots não são contíguos. Entretanto, uma requisição por um slot será alocada sem bloqueio.

Esse exemplo mostra dois problemas da abordagem desse algoritmo. O primeiro deles é a perda por fragmentação apresentada no estado (c). O segundo é que o algoritmo favorece a classe de menor ocupação, pois essa poderá preencher os “buracos” com mais facilidade que a outra classe. A esse favorecimento é dado o nome de parcialidade (*unfairness*) e à sua ausência o de imparcialidade (*fairness*).

2.2 Desfragmentação

Como introduzido anteriormente, na Seção 2.1.3, a desfragmentação é a técnica mais robusta no combate à fragmentação espectral. Ela consiste na “movimentação” das conexões já alocadas no espectro para evitar que novas conexões sejam bloqueadas por falta de blocos de slot contíguos durante o caminho encontrado.

2.2.1 Técnicas de desfragmentação

As técnicas de desfragmentação podem ser classificadas em dois grupos: reativas e proativas (CHATTERJEE *et al.*, 2015). Na primeira abordagem, a desfragmentação é executada quando o sistema alcança um limiar que deve ser definido pelo administrador do sistema. A

variável monitorada deve ser escolhida, assim como o seu valor de limiar. A escolha do valor impacta consideravelmente o desempenho do sistema dado, pois se for escolhido um valor muito comum, o sistema será interrompido constantemente, e se for um valor incomum, a desfragmentação não ocorrerá com frequência suficiente para melhorar o desempenho.

O segundo conjunto de técnicas, as reativas, seguem a lógica de sempre desfragmentar o sistema, assim que uma conexão sai do sistema. Mantendo o sistema sempre sob o estado desfragmentado, as perdas de fragmentação do sistema caem para zero. Esse tipo de desfragmentação é o escolhido para as análises presentes na seção de resultados (Seção 4.3).

Considerando mais a fundo as restrições de camada física, (CUGINI *et al.*, 2013) propõe uma técnica que não gera interrupção no tráfego (*hitless*), utilizando-se de reconfiguração de frequência de *lightpath* dinâmica. Além disso, essa técnica não necessita de *transponders* adicionais.

2.2.2 Alocação espectral e desfragmentação

Como mostrado anteriormente, o tráfego uniclasse alocado utilizando-se um algoritmo inteligente e definitivo, pode gerar fragmentação, mas somente o tráfego multiclasse pode gerar perda de fragmentação. Dessa forma, independentemente do objetivo do algoritmo de alocação, a desfragmentação se torna um requisito essencial para o funcionamento das redes ópticas elásticas multiclasse. Entretanto, como será mostrado adiante, nem todos os algoritmos de alocação multiclasse sofrem de perda de fragmentação.

Uma técnica de desfragmentação proativa pode ser vista em (ZHANG *et al.*, 2013), que descreve um algoritmo de RSA baseado em desfragmentação. Esse algoritmo tem como alvo reduzir a Probabilidade de Bloqueio (PB) geral do sistema, minimizando, simultaneamente, a quantidade de conexões movimentadas por operação de desfragmentação. Esse algoritmo sempre julga e movimenta somente 30% das conexões alcançando 1% de interrupção do sistema.

Outra técnica proativa é apresentada em (KADOHATA *et al.*, 2012). Aqui, é proposto que, uma vez por mês ou semana, seja feita uma “manutenção” do sistema. A operação periódica envolve o recálculo da topologia virtual da rede e a desfragmentação das conexões. Essas mudanças devem considerar a chegada e saída de conexões e alterações no tráfego das conexões já estabelecidas.

A Figura 5 mostra o funcionamento do algoritmo de alocação espectral de *First-Fit* utilizando a desfragmentação.

A principal diferença aqui é que, na transição do estado (b) para o estado (c), o sistema passa por uma desfragmentação proativa por causa da saída da conexão de um slot. Isso leva a todas as conexões serem reajustadas para a um índice a menos que o seu índice anterior. Dessa forma, se tornando um sistema com dois slots livres, assim como anteriormente (Figura 4.c),

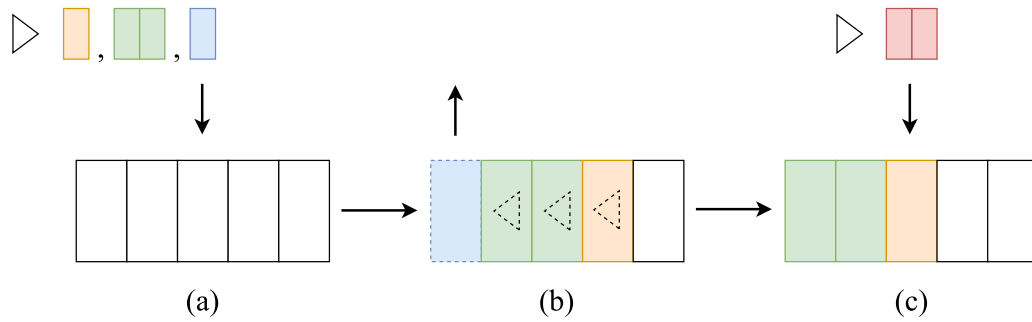


Figura 5 – Algoritmo de alocação First-Fit com desfragmentação.

entretanto, agora, esses slots são contíguos tornando o sistema aberto a conexões de dois slots ou um slot.

2.3 Imparcialidade e o seu custo

Dentro do escopo da operação de RMSA, os algoritmos desenvolvidos podem perseguir diferentes objetivos. Por exemplo, um algoritmo de RMSA pode objetivar manter a maior quantidade de clientes na rede simultaneamente. Outros podem desejar transferir a maior quantidade de dados possível por unidade de tempo.

Como apontado por (JAIN *et al.*, 1998), em sistemas computacionais que compartilham recursos, como o canal em sistemas de comunicações, a forma de distribuição desses recursos entre os seus usuários pode levar à discriminação de um ou mais indivíduos. Dessa forma, um meta-objetivo que pode ser perseguido por algoritmos de RMSA é a distribuição proporcional, ou imparcial, do recurso do canal entre suas classes de usuários.

No contexto das redes ópticas elásticas, os algoritmos de RMSA que otimizam o uso do espectro com a maior quantidade de usuários conectados tendem a dar prioridade a classes de usuários com menor ocupação (CALLEGATI *et al.*, 2013). Aumentado-se a quantidade de chegadas de requisições durante o tempo de vida médio de cada conexão, acontece um acúmulo de conexões no sistema. Para esse tipo de algoritmo, isso significa a presença quase exclusiva das menores classes. Essa discriminação pode significar um serviço pouco acessível para certos tipos de clientes, característica que alguns gerentes de sistemas podem não querer associado ao seu serviço.

A essa discriminação, dá-se o nome de parcialidade. Definir a medida do quão parcial é a distribuição do recurso não é uma tarefa simples. Esse valor pode ser medido de acordo com diversas métricas de desempenho, como, por exemplo, a porcentagem de uso do recurso por cada usuário. Dessa forma, (JAIN *et al.*, 1998) define um índice que pode ser aplicado a qualquer sistema de distribuição de recurso, contanto que se saiba o quanto cada usuário recebeu e quanto cada usuário pagou pelo seu acesso. Essa equação será explicada em mais detalhes a seguir, na Seção 2.4.4.

Com a imparcialidade como objetivo, diversos algoritmos de alocação espectral foram desenvolvidos no decorrer dos últimos anos. Esse trabalho discute duas dessas soluções: o particionamento do espectro e a política de reserva de entroncamento (TR), da sigla do nome em inglês (*Trunk Reservation*).

2.3.1 Particionamento

O particionamento do espectro é uma política de alocação que restringe o número máximo de conexões de cada classe e restringe o posicionamento dessas conexões a regiões pre-definidas para cada classe que não se sobrepõem. Exemplos de particionamento do espectro podem ser vistos na Figura 6 para um único enlace. No caso de uma rede, as partições tem que ser nas mesmas regiões em todos os enlaces.

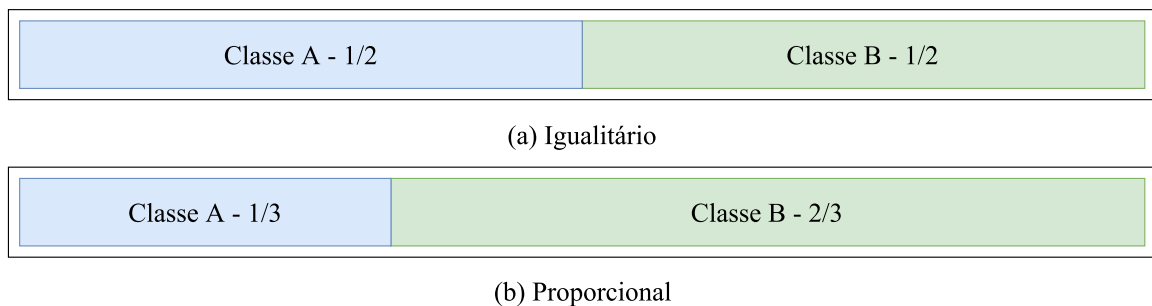


Figura 6 – Exemplos de particionamento do enlace óptico.

O particionamento mostrado na Figura 6.a é uma divisão do espectro em partes iguais para cada classe de acesso. Para tráfego uniforme, isto é, todas as classes de acesso requisitam conexões com a mesma frequência, esse particionamento pode ou não ser imparcial. Isso porque as classes de acesso podem ocupar o espectro de forma diferente com cada conexão. Assim, para esse caso, as classes com menor ocupação poderão alocar mais conexões simultaneamente, gerando um desbalanceio nas probabilidades de bloqueio.

O caso descrito na Figura 6.b é uma melhoria da solução anterior. Nessa solução, as partições são proporcionais à ocupação de cada classe e aos seus padrões de tráfego. Se for considerado que o tráfego é uniforme, as partições devem ser proporcionais somente aos tamanhos de cada conexão de cada classe. Na Figura, tem-se o particionamento para tráfego uniforme e classe B com conexões 2 vezes maiores que as conexões da classe A.

Apesar da simplicidade desse método, encontrar uma configuração de partições que gere imparcialidade pode ser um problema sem solução, dado que as razões entre as ocupações das conexões das classes de acesso (e/ou as taxas em seus perfis de tráfego) variam dentro do espaço real e os tamanhos das partições somente variam dentro do espaço natural (Figura 7). Assim, como pode ser visto na figura, somente algumas combinações de valores de partição geram parcialidade.

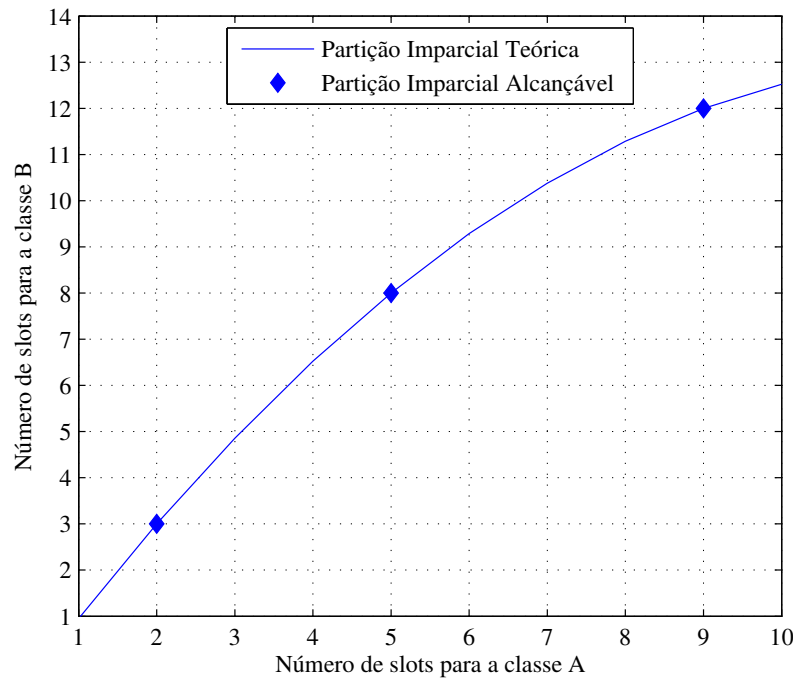


Figura 7 – Exemplo de espaço de soluções imparciais do particionamento para duas classes de acesso.

Entre o compartilhamento completo do espectro por todas as classes de acesso e a divisão completa do espectro em partições dedicadas ao uso exclusivo de cada classe, existe uma solução intermediária chamada de Pseudo-Particionamento (WANG; MUKHERJEE, 2014). Nessa política, o espectro ainda é dividido em partições, entretanto não mais de uso exclusivo de cada classe de acesso. Dessa forma, uma partição pode receber conexões de uma classe de acesso ou de um conjunto delas. Em comparação com o particionamento exclusivo do espectro e o compartilhamento completo, o Pseudo-Particionamento apresenta melhor desempenho em ocupação do espectro, para baixas cargas de tráfego.

Dessa forma, o Pseudo-Particionamento se apresenta como uma forma de melhorar o desempenho do Particionamento, quando a imparcialidade torna o seu desempenho muito inferior à opção convencional de alocação voraz.

2.3.2 Trunk Reservation

A política de alocação de Reserva de Entroncamento (TR) se caracteriza pela recusa de requisições caso o sistema esteja em um estado que discrimina alguma classe de acesso. A Figura 8 mostra exemplos de estados de enlaces e sua classificação quanto a favorecimento de classes específicas.

Considerando que existem duas classes de acesso, uma que requer um slot (A) e uma que requer dois slots (B), a figura mostra as 3 situações possíveis para o sistema. Na Figura 8.a,

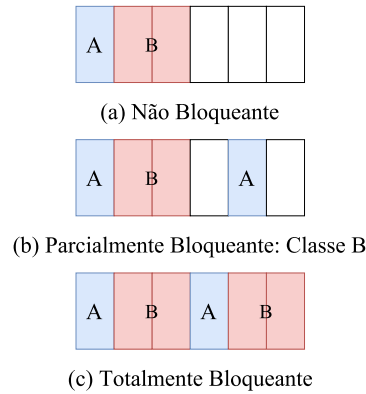


Figura 8 – Classificação de estados quanto ao favorecimento de classes.

o sistema pode alocar conexões de um ou dois slots, assim, esse tipo de estado é classificado como não bloqueante. Já na Figura 8.b, o enlace não aceita conexões da classe B porque a mesma requer dois slots contíguos. Por último, um exemplo de estado bloqueante para todas as classes é o estado mostrado na Figura 8.c. É importante notar que, dependendo do conjunto de ocupações de classes escolhido e do tamanho do espectro (em slots espectrais), os estados bloqueantes para todas as classes não necessariamente tem todos os slots ocupados, como ocorre para o exemplo da Figura 8.

Dessa forma, o que a política de alocação faz é impedir que seja alocada uma conexão, uma vez que o sistema está em um estado em que somente uma parcela das classes de acesso possa ser servida. Isso gera imparcialidade, dado que todos os estados que são bloqueantes para alguma classe, agora se tornam bloqueantes para todas, isto é, as probabilidades de bloqueio são iguais entre todas as classes de acesso. Ainda utilizando a Figura 8 como referência, o sistema irá aceitar requisições de todas as classes na situação 8.a e recusará todas as classes nos outros dois estados.

2.3.3 Custo da imparcialidade

O uso de políticas pode igualar as probabilidades de bloqueio e, assim, gerar imparcialidade. Entretanto, por recusar conexões em estados em que existe quantidade de slots disponíveis e contíguos, essas soluções geralmente levam a desempenhos inferiores quando comparadas sob a luz da eficiência na ocupação do espectro.

Em (WALDMAN *et al.*, 2016), foi mostrado que ambos o particionamento e o TR apresentam custo de utilização em termos de eficiência no número de slots ocupados na média em comparação com algoritmos vorazes como *First-Fit*.

Os resultados apresentados posteriormente nesse trabalho visam medir com mais detalhes esse *trade-off*, utilizando novas maneiras de medir essa diferença entre o desempenho dos dois tipos de solução.

2.4 Métricas de Avaliação de Desempenho

Os algoritmos de alocação espectral tem a função de alocar conexões no espectro, entretanto diferem entre si na forma como o fazem. Assim, alguns algoritmos podem ser mais adequados para certas situações em que outros podem apresentar desempenho inferior.

Dessa forma, para se poder classificar os algoritmos qualitativamente, a comparação entre esses algoritmos deve ser feita de forma sistemática por meio de métricas de desempenho. Essas métricas podem avaliar a capacidade do algoritmo em alocar muitas conexões simultaneamente, ou a porcentagem do espectro ocupado na média por ele, ou até se a distribuição do espectro entre os clientes da rede é justa. Cada métrica pode estar ligada a uma qualidade que se deseja avaliar em um certo conjunto de algoritmos.

Essa seção apresentará o conjunto de métricas utilizadas na seção de resultados para a avaliação dos algoritmos de alocação. Serão explicadas as informações qualitativas de cada uma, entretanto, os detalhes de cálculo desses valores serão explicados em detalhes somente no Capítulo 4.

2.4.1 Tráfego médio alcançado

O tráfego de informação é denominado aqui como a média temporal da quantidade de dados transferidos na rede. Considera-se sempre o melhor caso, no qual todos os clientes da rede estão transmitindo à maior taxa que podem para o espectro alocado para ele, ocupando 100% do seu bloco de slots designados.

Esse valor indica a eficiência espectral média do algoritmo. Quando comparados os valores de dois algoritmos, o com maior tráfego médio conseguiu alocar mais conexões que utilizam o espectro de forma mais eficiente.

2.4.2 Ocupação média

Um gerente pode considerar que a fibra é um recurso escasso que se deseja vender de forma mais eficiente possível. Assim, ele pode avaliar o desempenho de um algoritmo como proporcional à porção do espectro que ele mantém ocupado com clientes.

A ocupação média é uma média temporal da porcentagem do espectro ocupado, considerando que último é abstraído aqui como um conjunto de N slots espectrais de 12,5 GHz. Assim, a ocupação média é a divisão do número médio de slots ocupados pelo número total de slots do sistema.

2.4.3 Probabilidade de bloqueio

No contexto de redes de computadores, um bloqueio é uma operação de negação de alocação de uma nova conexão que foi requisitada por um cliente da rede. Bloqueios não são

ações bem vindas por gerentes pois indicam que o sistema está subdimensionado para o tráfego de requisições que está recebendo ou que não está alocando as conexões na rede de forma eficiente.

Assim, a forma de medir a frequência de bloqueios em uma rede é pela probabilidade de bloqueio. Esse valor é uma porcentagem que indica a chance de uma nova requisição ser recusada. No caso de tráfego multi-usuário, aqui estudado, o sistema pode apresentar valor diferente para cada classe. Assim, os resultados de probabilidade de bloqueio são apresentados por classe de acesso.

2.4.4 Índice de Imparcialidade

O termo *imparcialidade* deve ser entendido aqui como a habilidade de alocar conexões na rede mantendo uma cota do espectro proporcional ao tamanho da requisição de cada classe de acesso e à respectiva taxa de requisições. Isso se deve ao fato de que classes de acesso com menor ocupação espectral tem vantagem na alocação quando o sistema se encontra em estados com poucos slots sobrando, como exemplificado na Figura 3.c.

Surge, então, a necessidade de medir a discriminação entre as classes de acesso na alocação de conexões, ou, como utilizado aqui, a imparcialidade da mesma. Esse é um problema que se estende além do escopo de redes de computadores e tem soluções que foram propostas em outros contextos que podem ser, aqui, utilizadas.

O índice de imparcialidade aqui utilizado é o proposto por Jain (JAIN *et al.*, 1998) para medir a imparcialidade na alocação de recursos num sistema de computação distribuído. Esse índice pode ser descrito pela equação:

$$f(x) = \frac{[\sum_{i=1}^n x_i]^2}{n \sum_{i=1}^n x_i^2}, x_i \geq 0 \quad (2.1)$$

em que $f(x)$ é valor do índice de imparcialidade para a distribuição de recurso x , que pode ser entendido como um vetor de parcelas do recursos distribuído e x_i é a parcela do recurso alocado ao usuário i .

2.5 Fragmentação e enlace único

Como explicado na seção anterior, o bloqueio de requisições pode ter duas causas: a falta de recurso ou a fragmentação do espectro gerada pelas conexões já alocadas. A primeira causa deve ser evitada dimensionando a quantidade total de espectro oferecido para o padrão de tráfego esperado no sistema. Existem diversas formas de tratar o problema como, por exemplo, modulação por divisão espacial (SARIDIS *et al.*, 2015) e planejamento multi-período (STRAUB *et al.*, 2006) e incremental (VELASCO *et al.*, 2016).

A fragmentação do espectro ocorre com a saída de conexões, que podem deixar vacâncias no espectro entre grupos de conexões já alocadas (YIN *et al.*, 2013). No caso das redes legadas WDM, essa fragmentação não gera perdas pois todas as conexões ocupam a mesma parcela do espectro (tráfego de única classe), fazendo com que as vacâncias geradas sempre possam ser preenchidas por novas conexões. Com uso de multiplexação espacial (FIORANI *et al.*, 2016), a fragmentação podem acontecer entre modos ou fibras, tornando a alocação de conexões ainda mais complexa.

Entretanto, nas redes flexíveis, a fragmentação pode gerar bloqueios, dado que o padrão de tráfego pode conter demandas por conexões de ocupação espectral diferentes (tráfego multi-classe).

A fragmentação do espectro dentro de uma rede de fibras é chamada de fragmentação horizontal e está exemplificada na Figura 9. Nela, as conexões já estabelecidas impedem que seja criada uma nova conexão entre os pontos A e C, pois essa só seria possível caso não houvesse a restrição de continuidade.

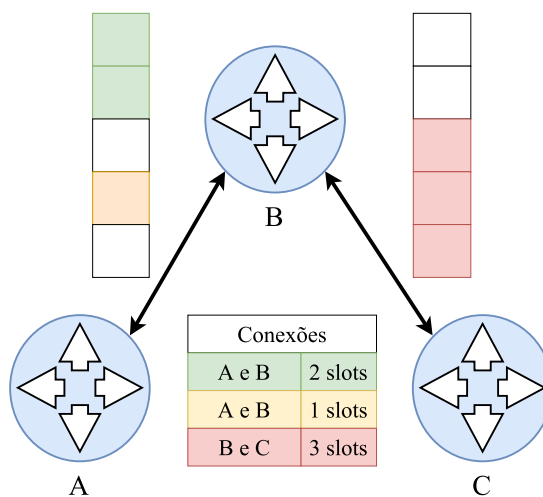
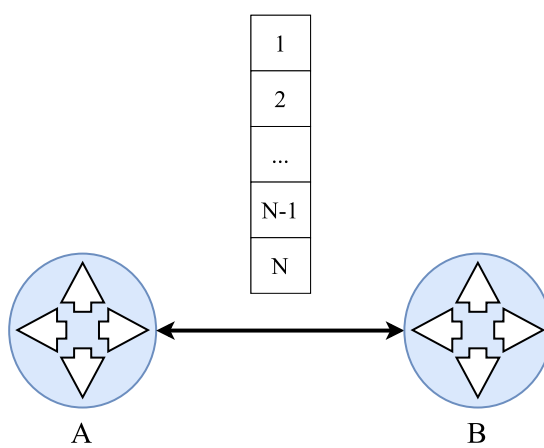


Figura 9 – Exemplos de fragmentação espectral em redes elásticas.

O outro tipo de fragmentação ocorre dentro de uma única fibra e é chamado de fragmentação vertical. Ainda na Figura 9, é possível observar, no enlace entre A e B, que existem dois *slots* espectrais não utilizados, entretanto por não serem contíguos um ao outro, não podem abrigar conexões que requerem dois *slots*, somente duas conexões de um *slot*, cada.

Dessa forma, para que os efeitos da fragmentação vertical sejam isolados, é preciso que a rede siga uma topologia de enlace único (Figura 10), ou seja, só tenha dois elementos multiplexadores ligados por uma única fibra. Cada conexão será bidirecional e ocupará uma quantidade de *slots* igual à definida para sua classe de acesso. Esse valores podem variar de 1 a N.

Figura 10 – Modelo de enlace único con N slots.

Parte II

Contribuições

3 Modelo analítico

As redes ópticas elásticas se configuram como redes de comutação de circuitos. Isso porque quando um usuário requisita uma conexão dentro da rede, os recursos da mesma devem ser alocados exclusivamente durante todo o período de vida da conexão solicitada.

Por causa disso, os desempenhos dos algoritmos de alocação estudados nesse trabalho foram obtidas analiticamente aproveitando-se do fato de que redes de comutação de circuitos podem ser modeladas utilizando teoria de filas para se obter as métricas de desempenho desejadas.

Esse capítulo se propõe a trazer uma base conceitual da modelagem da rede como um sistema de filas. A Seção 3.1 contém toda a modelagem matemática do sistema, explicando cada passo e consideração feita para obtenção do conjunto de equações utilizadas posteriormente para coleta das métricas. A Seção 3.2 explica a importância da técnica de desfragmentação espectral para a modelagem obtida do sistema. Em seguida, a Seção 3.3 adiciona o impacto das políticas de alocação imparciais no modelo obtido. Por fim, são mostradas as limitações dessa metodologia na Seção 3.4.

3.1 Modelagem do sistema

Essa seção é uma breve síntese dos principais passos para a modelagem do sistema. Conhecimentos básicos de teoria de filas são necessários para o total entendimento dos teoremas demonstrados aqui. As demonstrações aqui explicadas seguem descritas em mais detalhes em (KUMAR *et al.*, 2004a).

O tráfego de dados indica a medida da quantidade de informação que está sendo transferida na rede. Ele pode ser medido pela carga de tráfego oferecida, que pode ser definida por um processo de chegada de requisições por conexões e uma distribuição da duração dessas conexões. A taxa de chegada de requisições pode ser definida por:

$$\lambda := \lim_{t \rightarrow \infty} (\text{Número de Chegadas em } [0, t]) / t$$

A carga de tráfego oferecida, mencionada anteriormente, pode ser equacionada da seguinte forma:

$$\rho := \frac{t_v}{t_c} = \frac{\lambda}{\mu} \quad (3.1)$$

Em que t_v é o tempo médio de vida de uma chamada (ou o inverso da taxa de término de

conexões μ) e t_c é o tempo médio entre chegadas de conexões (ou o inverso da taxa de chegada de requisições λ). O valor da carga de tráfego não tem dimensão, mas sua unidade é chamada erlang. Esse valor pode ser entendido como o número médio de chegadas durante o tempo de vida de uma conexão ou o número médio de conexões ativas no espectro, se a sua capacidade fosse infinita.

No contexto de criação de um modelo de negócio, uma classe de acesso pode ser definida como um conjunto de usuários com necessidades semelhantes, dos quais serão cobrados valores semelhantes pelo uso do espectro. Entretanto, na modelagem do sistema, uma classe de acesso é definida pelo conjunto: taxa de chegada de requisições, tempo de vida médio de uma conexão e recurso ocupado por uma conexão.

Quando não há mais espectro disponível para alocar uma requisição, ela é recusada. Uma requisição recusada pelo sistema não retorna à fila de requisições, isto é, ela é perdida e não pode mais ser recuperada. Por causa disso, esse tipo de sistema é considerado um modelo de perdas e sua principal métrica de performance é a probabilidade de bloqueio de requisições (probabilidade das conexões serem recusadas). A probabilidade de bloqueio também pode ser utilizada para se obter a carga efetivamente carregada pelo sistema (via Teorema de Little (LITTLE; GRAVES, 2008)).

Considerando C o número de slots espectrais de um enlace, esse valor define o tamanho total de slots dos sistema, dado que a topologia estudada é a de enlace único. Existem K classes de acesso. Cada classe k , $k \in \{1, 2, \dots, K\}$, tem um processo estocástico de chegada de requisições Poissoniano estacionário de taxa λ_k . As conexões tem tempos de vida independentes e identicamente distribuídos segundo uma distribuição exponencial de média μ_k^{-1} e requerem b_k slots espectrais quando ativa, sendo k o índice da classe à qual pertence a conexão.

Inicialmente, será modelado um sistema sob algoritmo de alocação voraz, que servirá de base para a apresentação dos modelos sob os outros algoritmos. Para esse sistema, o requisito para aceitação de uma nova requisição da classe k é que existam b_k slots espectrais contíguos disponíveis. Entretanto, como está sendo pressuposto o uso de desfragmentação espectral, sempre que existirem b_k slots espectrais disponíveis, os mesmos serão contíguos.

O estado do enlace num momento t qualquer pode ser descrito pelo conjunto dos números de conexões ativas n_k de cada classe de acesso k . Assim, definem-se os dois vetores coluna a seguir:

$$\vec{b} := [b_1, \dots, b_K]^T, \quad b_k \in \mathbb{N}$$

$$\vec{n} := [n_1, \dots, n_K]^T, \quad n_k \in \mathbb{N}$$

em que \vec{b} é o vetor que descreve o número de slots por conexão de cada classe e \vec{n} é

o vetor que descreve o número de conexões ativas de cada classe em um dado instante. A partir dessa notação, desconsiderando inicialmente a restrição de contiguidade, pode-se definir a regra geral de aceitabilidade de requisições da classe k' como:

$$b_{k'} \leq C - \vec{n}^T \cdot \vec{b}$$

Dessa forma, dependendo do estado em que o sistema se encontra, uma requisição pode ou não ser bloqueada. Como as chegadas são consideradas Poissonianas e chegadas Poissonianas veem médias temporais (a propriedade PASTA, explicada no apêndice D do livro de referência (KUMAR *et al.*, 2004b)), a probabilidade de uma chegada ter que ser recusada é a probabilidade de o sistema estar em um estado bloqueante. Assim, encontrando-se as probabilidades estacionárias de estado, diversas métricas podem ser calculadas, como as probabilidades de bloqueio ou ocupação média do espectro.

Considerando $X_k(t)$ um inteiro que descreve o número de conexões ativas da classe k num tempo t , tem-se o vetor:

$$\vec{X}(t) := [X_1(t), \dots, X_k(t), \dots, X_K(t)]^T, t \geq 0$$

como descrição do processo estocástico do número de conexões no enlace em um dado instante de tempo t . A princípio, será considerado que não será imposto limite na quantidade de conexões que cada classe de acesso pode inserir, ou seja, o espectro tem infinitos slots.

Assim, porque as chegadas são poissonianas e os tempos de vida das conexões são distribuídos exponencialmente e independentes e identicamente distribuídos, o processo $\{\vec{X}(t), t \geq 0\}$ é uma cadeia de Markov de tempo contínuo ou CTMC (do inglês *Continuous Time Markov Chain*). Como foi considerado que modelo é um sistema de perdas e todos os clientes são servidos imediatamente, a taxa de término de conexões da classe k quando $\vec{X}(t) = \vec{n} = [n_1, \dots, n_k, \dots, n_K]$ é $n_k \mu_k$. A Figura 11 mostra uma CTMC que descreve o comportamento de um enlace com infinitos slots que carrega um tráfego com duas classes de acesso, um que ocupa 2 e outro que ocupa 4 slots.

Pode ser demonstrado que essas CTMCs são irredutíveis e recorrente positivas, entretanto essas demonstrações não feitas nesse trabalho, mais detalhes podem ser encontrados no apêndice D do livro de referência (KUMAR *et al.*, 2004b). Essas propriedades garantem que as cadeias possuem distribuição de probabilidade estacionária.

Outra propriedade que pode ser observada nas cadeias descritas anteriormente é a reversibilidade. Segundo o critério de reversibilidade de Kolmogorov, uma CTMC irredutível e recorrente positiva é reversível se, para qualquer percurso de estados fechado, o produto das taxas de transição em um sentido for igual ao produto das taxas de transição no sentido oposto.

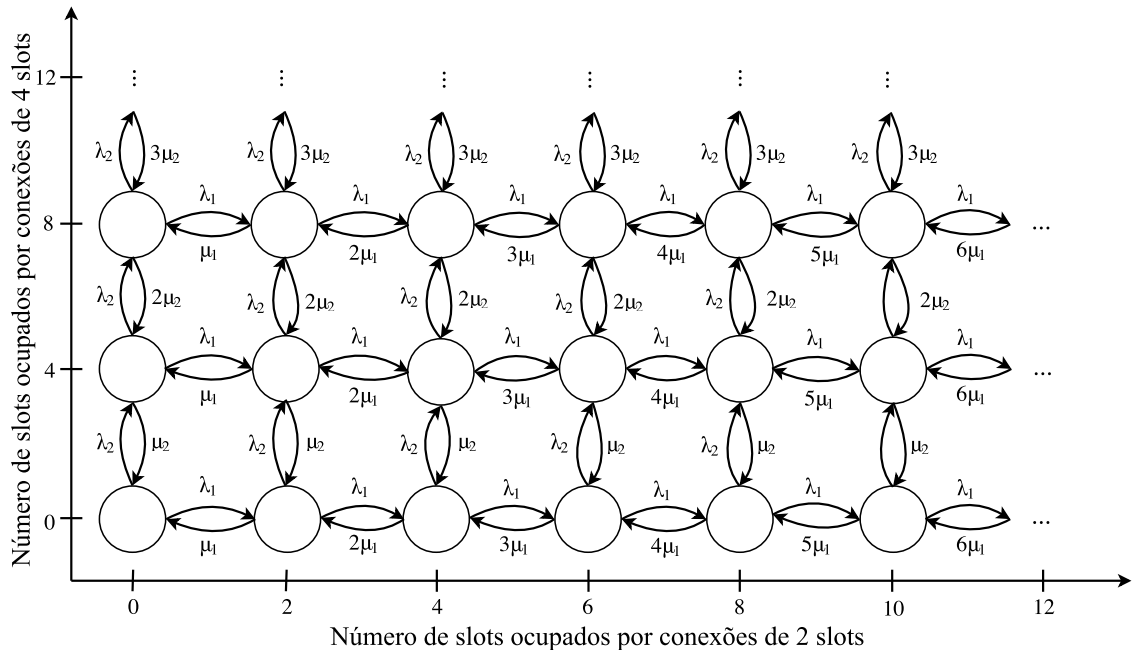


Figura 11 – A CTMC para um algoritmo voraz desfragmentado com $\vec{b} = [2, 4]^T$ e $C = \infty$.

Será considerada a seguinte notação de percurso: $\{(A,B) \rightarrow (C,D)\}$, como sendo a transição entre o estado com A conexões de 2 slots e B conexões de 4 slots para o estado de C de conexões de 2 slots e D conexões de 4 slots. Dessa forma, utilizando a cadeia da Figura 11 como referência, caso seja escolhido o percurso $\{(0,0) \rightarrow (0,1) \rightarrow (1,1) \rightarrow (1,0) \rightarrow (0,0)\}$, o produto das taxas de transição no sentido anti-horário é $\lambda_1 \lambda_2 \mu_1 \mu_2$. O mesmo produto é obtido quando calculado no sentido horário. O mesmo pode ser observado para qualquer outro percurso fechado desta cadeia.

Assim, cada classe pode ser descrita individualmente por uma CTMC em árvore e, por isso, e equações de balanceamento podem ser escritas como:

$$\lambda_k \pi_\infty(n_k) = (n_k + 1) \mu_k \pi_\infty(n_k + 1), \quad n_k \geq 0 \quad (3.2)$$

Em que $\pi_\infty(n_k)$ é a probabilidade de o sistema ter n_k conexões da classe k , dado que está operando sem limites de alocação. Resolvendo essas equações, tem-se:

$$\pi_\infty(n_k) = e^{-\rho_k} \frac{\rho_k^{n_k}}{n_k!} \quad (3.3)$$

Em que ρ_k é a carga de tráfego oferecida pela classe de acesso k (definido na Equação 3.1). Assim, a probabilidade estacionária de que o sistema esteja no estado $\vec{X}(t) = \vec{n} = [n_1, \dots, n_K]$, ou seja a probabilidade $\pi_\infty(\vec{n})$, é o produto das probabilidades individuais de cada

elemento do vetor, ou seja:

$$\pi_\infty(\vec{n}) := \pi_\infty(n_1, \dots, n_K) = \prod_{k=1}^K \pi_\infty(n_k) = \prod_{k=1}^K e^{-\rho_k} \frac{\rho_k^{n_k}}{n_k!} = e^{-\sum_k \rho_k} \prod_{k=1}^K \frac{\rho_k^{n_k}}{n_k!} \quad (3.4)$$

Considerando \vec{e}_k um vetor em que o k -ésimo elemento é 1 e outros elementos são 0, o estado $\vec{n} + \vec{e}_k$ terá uma conexão a mais da classe k do que o estado \vec{n} e o estado $\vec{n} - \vec{e}_k$, terá uma conexão a menos. Assim, para $k = 1, \dots, K$, a relação entre os estados \vec{n} e o estado $\vec{n} + \vec{e}_k$ é dada por:

$$\begin{aligned} \pi_\infty(\vec{n} + \vec{e}_k) &= \pi_\infty(n_k + 1) \prod_{\substack{k'=1 \\ k' \neq k}}^K \pi_\infty(n_{k'}) \\ &= \frac{\lambda_k}{(n_k + 1)\mu_k} \pi_\infty(n_k) \prod_{\substack{k'=1 \\ k' \neq k}}^K \pi_\infty(n_{k'}) \\ \lambda_k \pi_\infty(\vec{n}) &= (n_k + 1)\mu_k \pi_\infty(\vec{n} + \vec{e}_k) \end{aligned} \quad (3.5)$$

Em que a segunda linha utiliza a equação de balanceamento detalhada de $X_k(t)$ (Equação 3.2). Como as únicas transições da CTMC $\vec{X}(t)$ são $\vec{n} + \vec{e}_k$ ou $\vec{n} - \vec{e}_k$ e o sistema obedece a uma equação de balanceamento, a cadeia é reversível, o que também pode ser observado traçando qualquer caminho fechado e testando o critério de Kolmogorov.

Entretanto, as cadeias formadas do funcionamento das redes elásticas não podem ser descritas por esse tipo de cadeia mostrado anteriormente. Isso porque o número de conexões que cada classe pode alocar é limitado pelo número de slots do espectro e pelas conexões já alocadas. Por exemplo, a Figura 12 mostra a CTMC para um espectro composto por 12 slots ($C = 12$), com classes de acesso $\vec{b} = [2, 4]^T$.

A limitação do número de conexões cria um truncamento na CTMC. Esse truncamento preserva a reversibilidade da CTMC, de acordo com o Teorema D.22 da Seção D.3.1 do livro (KUMAR *et al.*, 2004a). Isso implica que as probabilidades estacionárias podem, então, serem obtidas a partir de:

$$\pi_C(n_1, n_2, \dots, n_K) = \frac{\pi_\infty(n_1, n_2, \dots, n_K)}{\sum_{\vec{n} \in S} \pi_\infty(n_1, n_2, \dots, n_K)} = \frac{e^{-\sum_k \rho_k} \prod_{k=1}^K \frac{\rho_k^{n_k}}{n_k!}}{\sum_{\vec{n} \in S} (e^{-\sum_k \rho_k} \prod_{k=1}^K \frac{\rho_k^{n_k}}{n_k!})} \quad (3.6)$$

Em que S é o conjunto de todos os possíveis estado nos quais o sistema pode estar. Dessa forma, considerando que o termo $e^{-\sum_k \rho_k}$ pode ser posto em evidência no denominador

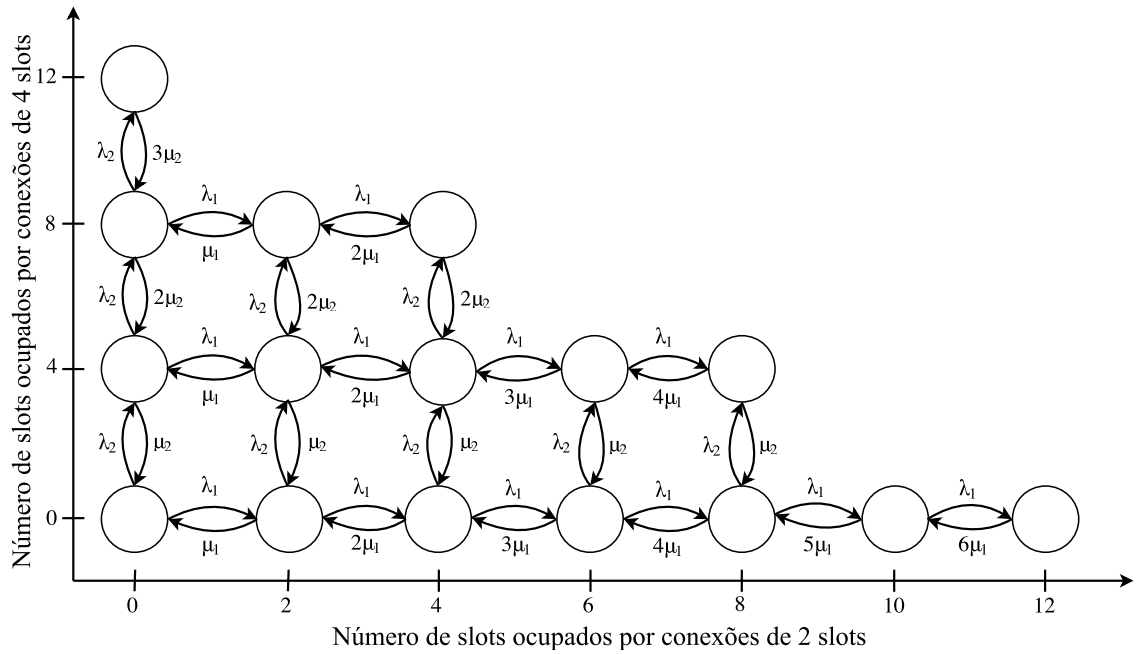


Figura 12 – A CTMC para um algoritmo voraz desfragmentado com $\vec{b} = [2, 4]^T$ e $C = 12$.

e, assim, ser removido da equação, a probabilidade descrita na Equação 3.6 pode ser reescrita como:

$$\pi_C(n_1, n_2, \dots, n_K) = \frac{1}{G} \prod_{k=1}^K \frac{\rho_k^{n_k}}{n_k!} \quad (3.7)$$

Em que a constante G é o fator normalizador descrito por:

$$G := \sum_{\vec{n} \in S} \prod_{k=1}^K \frac{\rho_k^{n_k}}{n_k!} \quad (3.8)$$

As equações 3.7 e 3.8 podem ser utilizadas para se obter as probabilidades de estado estacionárias de qualquer CTMC gerada pelos algoritmos de alocação estudados nesse trabalho, contanto que seja reversível. Uma vantagem dessa metodologia é que a complexidade computacional cresce somente com quadrado do número de slots.

3.2 Desfragmentação e Reversibilidade

Na seção anterior, foi feita a consideração inicial de que as classes de acesso não tinham limite de quantidade máxima de conexões simultâneas. Isso levou a cadeias em árvore para cada classe de acesso. No contexto dos algoritmos de alocação de recursos finitos, as CTMC obtidas não obedecem esse conceito pois a limitação no número de conexões simultâneas e

as ocupações das outras classe podem interferir na aceitação de uma conexão de uma classe específica.

O algoritmo *First-Fit* é um algoritmo de alocação espectral considerado voraz. Sua regra de alocação se resume a alocar uma conexão sempre que houver uma quantidade contígua de slots não utilizados, dando preferência sempre às frequências com menor valor. A Figura 13, extraída de (WALDMAN *et al.*, 2015), descreve a CTMC para esse algoritmo para $C = 6$ e $\vec{b} = [2, 4]^T$. Os estados simétricos estão agrupados por equivalência para uma visualização simplificada do comportamento do sistema. Dessa forma, as letras que definem o estado são a posição de cada conexão, assim “e” é uma conexão à esquerda do espectro (frequências mais baixas) e “eE” são duas conexões à esquerda. Se uma letra estiver em maiúsculo, ela é uma conexão de 4 slots, se não, de 2 slots.

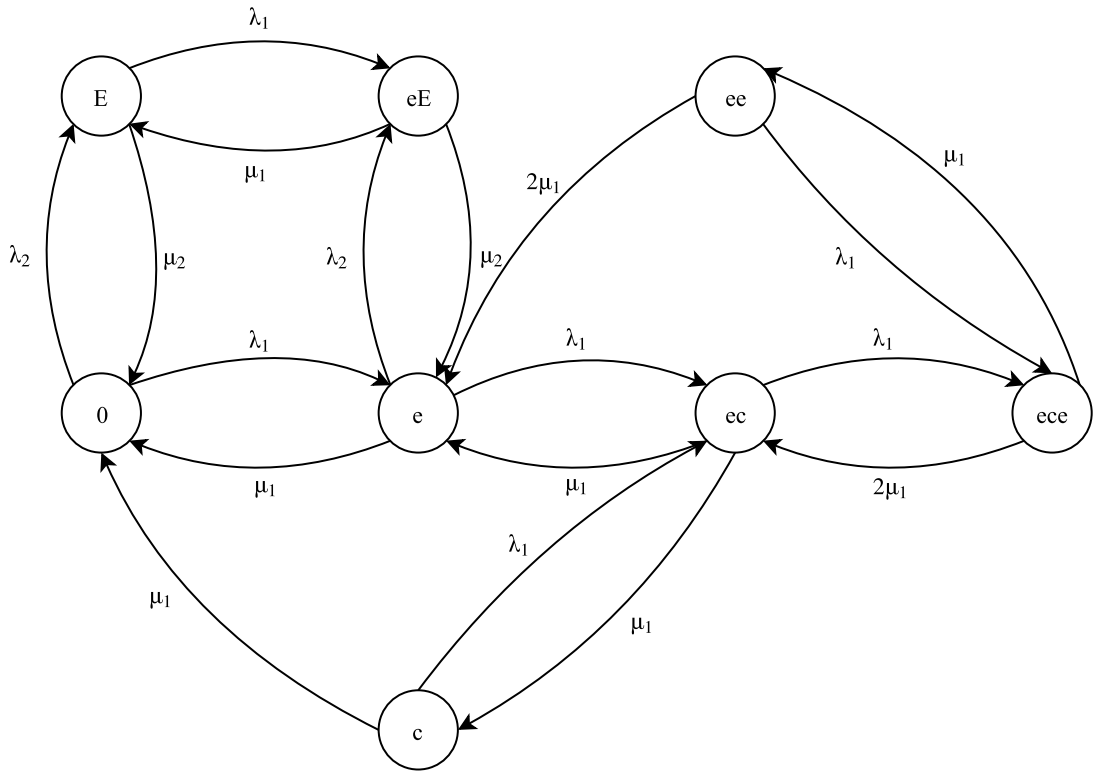


Figura 13 – A CTMC reduzida para o algoritmo *First-Fit* com $\vec{b} = [2, 4]^T$ e $C = 6$.

Aplicando a técnica de desfragmentação, a CTMC obtida pode ser observada na Figura 14.

A CTMC não desfragmentada (Figura 13) apresenta diversas diferenças em relação à sua versão desfragmentada. Uma dessas diferenças é a assimetria da primeira cadeia, advinda do fato de que os estados são unicamente definidos não só pelo número de conexões de cada classe, mas também pela posição de cada uma. Isso leva uma explosão combinatória no número de estados com o aumento de C , dificultando o cálculo analítico com o aumento do tamanho do espectro no modelo.

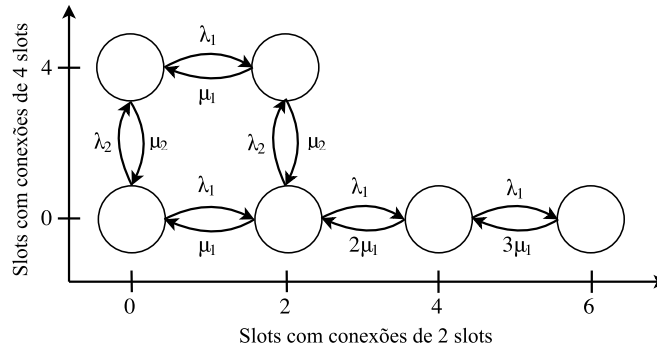


Figura 14 – A CTMC para o algoritmo *First-Fit* desfragmentado com $\vec{b} = [2, 4]^T$ e $C = 6$.

Outra diferença observada é a reversibilidade. Assim como as cadeias analisadas na Seção 3.1, as CTMCs para algoritmos de alocação que utilizam desfragmentação tem estados unicamente definidos pelo número de conexões de cada classe e apresentam reversibilidade (observável pelo critério de Kolmogorov). Entretanto, sem a desfragmentação, não há garantia de reversibilidade, o que pode ser constatado no exemplo da Figura 13. Nesse caso, as transições unidirecionais entre os pares de estados $\{ee, e\}$ e $\{c, 0\}$ fazem com que o critério de Kolmogorov não seja cumprido, demonstrando sua falta de reversibilidade.

Dessa forma, como a desfragmentação foi utilizada em todos os cenários, a maior parte dos casos estudados pode ser avaliada a partir das probabilidades de estado obtidas utilizando as equações 3.7 e 3.8. Entretanto, existem casos de algoritmo não-vorazes nos quais, mesmo com a desfragmentação, o algoritmo de alocação não gera uma CTMC reversível. A Seção 3.3.2 irá definir quais são esses casos e como calcular as probabilidades de estado estacionárias em CTMCs não-reversíveis.

3.3 Algoritmos baseados em políticas não-vorazes

Nos algoritmos vorazes, as requisições de conexão só são recusadas quando não há slots contíguos no espectro para comportar essas requisições. O mesmo não pode ser dito para a outra categoria de algoritmos que serão analisados nesse trabalho: os algoritmos baseados em políticas imparciais.

Nesse tipo de algoritmo, as requisições podem ser recusadas por outras razões, levando a CTMCs diferentes das observadas para algoritmos vorazes. Dessa forma, essa seção irá descrever as diferenças nas CTMCs geradas a partir de duas políticas imparciais: o *Particionamento* e o *Trunk Reservation*.

3.3.1 Particionamento

Existem diversas formas de particionar o espectro. No contexto desse trabalho, o particionamento sempre irá buscar a imparcialidade. Outra consideração importante é que as conexões dentro de cada partição serão alocadas utilizando *First-Fit* sobre uma única classe, impedindo que ocorra perda por fragmentação dentro de cada partição. As partições serão descritas nesse trabalho com o vetor:

$$\vec{p} = [p_1, p_2, \dots, p_P]$$

Em que cada valor p_k do vetor descreve o número de slots reservados para aquela partição e P é o número de partições. Assim, para o particionamento mutuamente exclusivo, em que não existem regiões compartilhadas no espectro, a regra para criação das partições é:

$$\sum_{k=1}^P p_k = C$$

Mantidas as mesmas condições de tráfego e banda, uma cadeia de Markov do particionamento será um truncamento da cadeia de Markov do algoritmo voraz. A Figura 15 mostra um exemplo do truncamento ocorrido quando particionado o espectro.

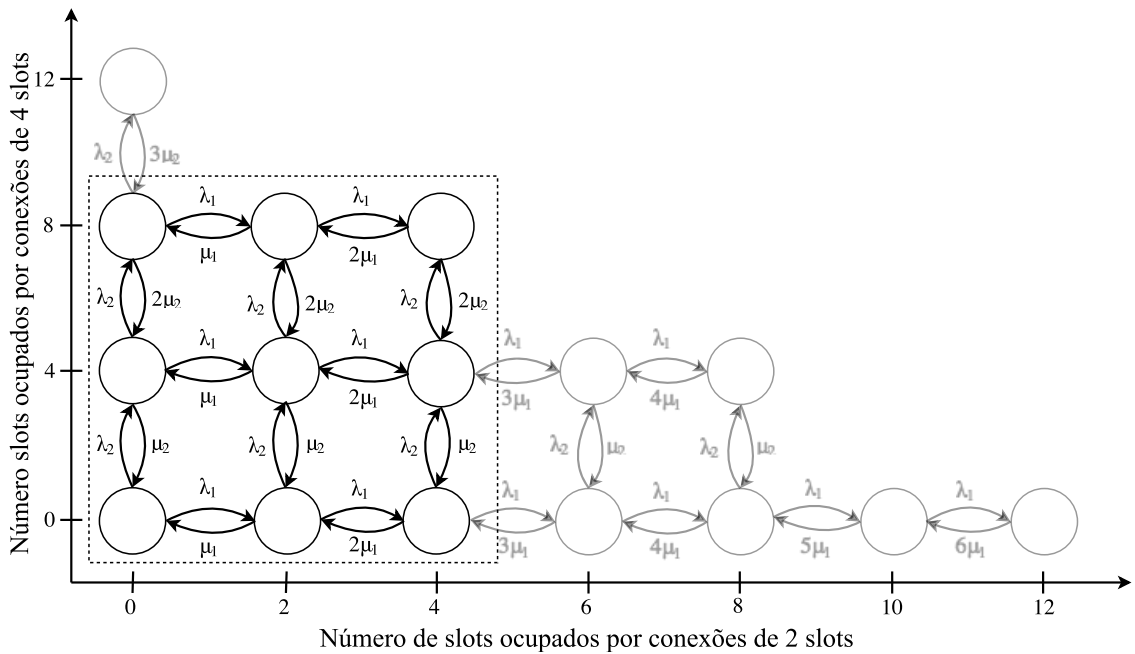


Figura 15 – Comparação das CTMCs entre o algoritmo voraz (todos os estados) e o algoritmo de Particionamento (estados dentro do retângulo) com $\vec{b} = [2, 4]^T$, $C = 12$ e $\vec{p} = [4, 8]$.

Desse truncamento, é importante destacar duas propriedades da CTMC resultante. A primeira delas é que o número de conexões de uma classe não interfere na alocação de conexões para a outra classe, somente o tamanho de cada partição impede a alocação de novas conexões. A segunda propriedade é que a cadeia continua cumprindo o critério de Kolmogorov e, por isso, as equações 3.7 e 3.8 podem ser utilizadas para se obter as probabilidades de estado estacionárias.

Outro tipo de particionamento explorado nesse trabalho é o pseudo-particionamento. Nesse tipo de particionamento, as partições podem não ser de uso exclusivo de uma classe, gerando a possibilidade de serem criadas partições para grupos de classe com ocupações semelhantes ou que não geram fragmentação.

Com o uso da desfragmentação, ou a escolha das classes de uma partição que não gere fragmentação, uma CTMC reversível, semelhante às obtidas anteriormente para os algoritmos vorazes e para o particionamento simples, pode ser obtida. O exemplo da Figura 16 mostra uma cadeia proveniente do particionamento de um espectro de 12 slots em três partições, mantidas as mesmas classes de acesso dos exemplos anteriores: p_1 e p_2 são partições de acesso privativo das classes de ocupação 2 e 4 e p_3 é uma partição de uso compartilhado.

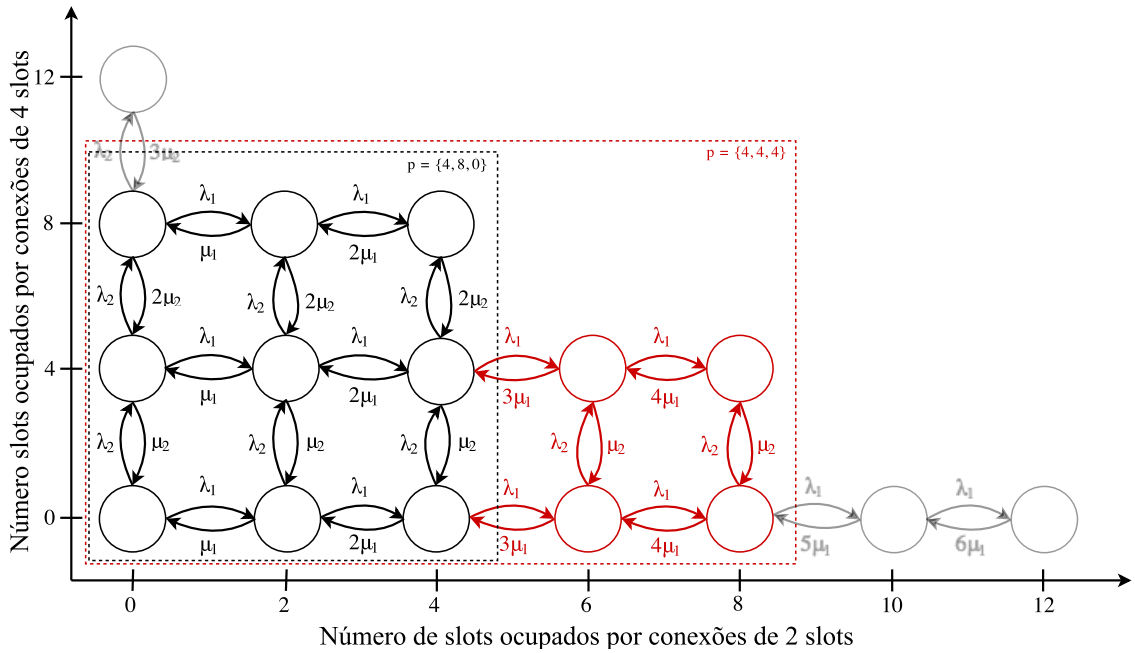


Figura 16 – Construção da CTMC do Pseudo-particionamento (com desfragmentação) a partir da CTMC do algoritmo voraz desfragmentado com $\vec{b} = [2, 4]^T$, $C = 12$.

É importante destacar que a CTMC da Figura 16 só mantém a reversibilidade porque as classes tem uma relação que impede fragmentação na partição de uso compartilhado e também permite a agregação de estados semelhantes. Além disso, com apenas duas classes, e usando *First-Fit* para uma e *Last-Fit* para outra, com desfragmentação para ambas, é possível manter

a contiguidade de todos os slots disponíveis, preservando a reversibilidade da CTMC. Outra característica dessas novas cadeias é a interferência do número de conexões de uma classe na possibilidade de alocação da outra. Isso ocorre por causa da partição compartilhada.

3.3.2 Trunk Reservation

O algoritmo de Reserva de Entroncamento (*Trunk Reservation*) altera a CTMC de uma forma diferente da observada para o particionamento. A lógica por trás da operação dessa política é a supressão das transições de estados que levem ao favorecimento de uma classe sobre outra. Numa CTMC, esses estados podem ser identificados pela falta de transições para uma das direções dos eixos que representam cada classe de acesso. Assim, a política de TR altera a cadeia suprimindo algumas transições e até estados.

O exemplo das Figuras 17 e 18 ilustra essa transformação na cadeia. Nessas figuras, todos os estados bloqueantes são identificados por um texto e coloração diferenciada. O texto corresponde à(s) classe(s) de acesso que aquele estado bloqueia, nesse caso, ele pode ser ‘2’, ‘4’ ou ‘TODAS’, quando o estado não aceita mais conexões (estado completo).

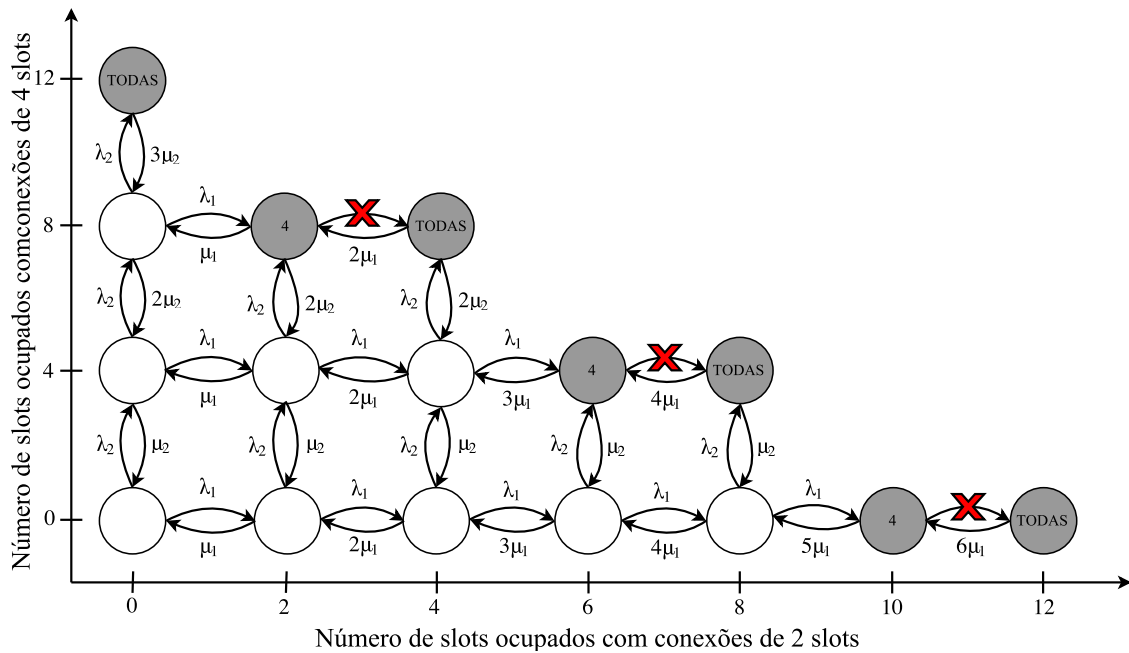


Figura 17 – Construção da CTMC do *Trunk Reservation* (com desfragmentação) a partir da CTMC do algoritmo voraz desfragmentado com $\vec{b} = [2, 4]^T$, $C = 12$.

Criada a partir da CTMC dos algoritmos vorazes com uso de desfragmentação, a Figura 17 explicita a remoção das transições destacadas com um ‘X’. Assim, a CTMC se torna a observada na Figura 18, pois os estados que bloqueavam somente a classe de acesso de 4 slots agora bloqueiam todas as classes de acesso e o estado (6,0) deixa de existir pois se torna inalcançável.

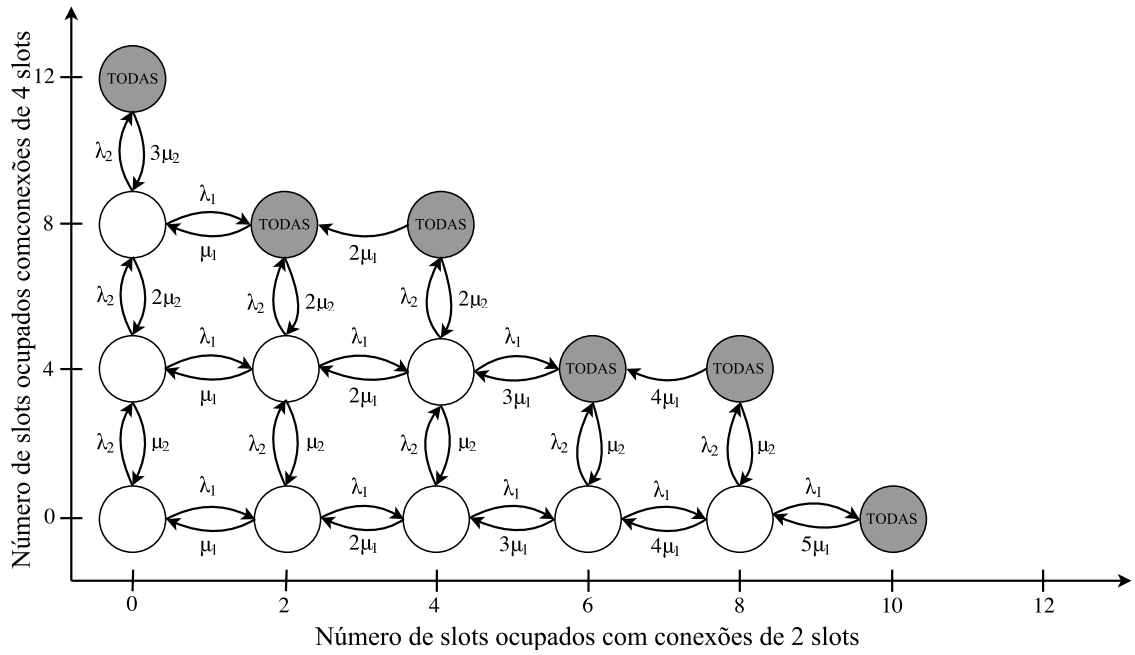


Figura 18 – CTMC do *Trunk Reservation* (com desfragmentação) para $\vec{b} = [2, 4]^T$, $C = 12$.

Embora a obtenção da CTMC para esse algoritmo seja sistemática, ela leva a uma CTMC que não é reversível, como pode ser observado pelo não cumprimento do critério de reversibilidade de Kolmogorov. Assim, para o cálculo desse tipo de cadeia, devem ser utilizadas somente as equações de balanceamento globais.

Para esse caso, o restante dessa seção desenvolve uma metodologia analítica para obtenção das probabilidades de estado.

Para uma CTMC qualquer, para todo estado i , no equilíbrio, tem-se que:

$$\pi_i \times \sum_j q_{i,j} = \sum_j \pi_j \times q_{j,i} \quad (3.9)$$

que π_i é a probabilidade de o sistema estar no estado i e $q_{i,j}$ é a taxa de transição do estado i para o estado j , onde se convencionou que $q_{i,i} = 0$ para todo i .

Assim, considerando a cadeia tem σ estados, tem-se o vetor de probabilidades de estado

$$\vec{\pi} = [\pi_1 \quad \pi_2 \quad \cdots \quad \pi_\sigma]$$

e a matriz de transição de estados

$$\mathbf{M} = \begin{bmatrix} q_{1,1} & q_{1,2} & \cdots & q_{1,j} & \cdots & q_{1,\sigma} \\ q_{2,1} & q_{2,2} & \cdots & q_{2,j} & \cdots & q_{2,\sigma} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ q_{\sigma,1} & q_{\sigma,2} & \cdots & q_{\sigma,j} & \cdots & q_{\sigma,\sigma} \end{bmatrix}.$$

Dessa forma, pode ser observado que o produto $\vec{\pi} \cdot \mathbf{M}$ leva a::

$$\begin{aligned} \vec{\pi} \cdot \mathbf{M} &= \begin{bmatrix} \sum_j \pi_j q_{j,1} & \sum_j \pi_j q_{j,2} & \cdots & \sum_j \pi_j q_{j,\sigma} \end{bmatrix} \\ &\stackrel{\text{eq.}}{=} \begin{bmatrix} \pi_1 \sum_j q_{1,j} & \pi_2 \sum_j q_{2,j} & \cdots & \pi_\sigma \sum_j q_{\sigma,j} \end{bmatrix} \\ &= \vec{\pi} \cdot \begin{bmatrix} \sum_j q_{1,j} & 0 & \cdots & 0 \\ 0 & \sum_j q_{2,j} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sum_j q_{\sigma,j} \end{bmatrix} \end{aligned} \quad (3.10)$$

A segunda igualdade na Equação 3.10 vem da propriedade explicitada na Equação 3.9.

Dessa forma, reescrevendo a matriz \mathbf{M} na igualdade descrita na Equação 3.10 e com certa manipulação algébrica, pode-se notar que:

$$\vec{\pi} \left\{ \begin{bmatrix} 0 & q_{1,2} & \cdots & q_{1,\sigma} \\ q_{2,1} & 0 & \cdots & q_{2,\sigma} \\ \vdots & \vdots & \ddots & \vdots \\ q_{\sigma,1} & q_{\sigma,2} & \cdots & 0 \end{bmatrix} - \begin{bmatrix} \sum_j q_{1,j} & 0 & \cdots & 0 \\ 0 & \sum_j q_{2,j} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sum_j q_{\sigma,j} \end{bmatrix} \right\} = 0 \quad (3.11)$$

Nomeando a matriz resultante da operação entre chaves na Equação 3.11 de \mathbf{Q} , tem-se:

$$\mathbf{Q} = \begin{bmatrix} -\sum_j q_{1,j} & q_{1,2} & \cdots & q_{1,\sigma} \\ q_{2,1} & -\sum_j q_{2,j} & \cdots & q_{2,\sigma} \\ \vdots & \vdots & \ddots & \vdots \\ q_{\sigma,1} & q_{\sigma,2} & \cdots & -\sum_j q_{\sigma,j} \end{bmatrix}$$

Ou seja, cada elemento da diagonal principal é o negativo da soma dos elementos de cada linha. Assim, dado que as taxas de transição de estado podem ser obtidas sistematicamente, foi obtido um sistema de equações lineares para o cálculo das probabilidades de estado $\vec{\pi}$ descrito por:

$$\vec{\pi} \cdot \mathbf{Q} = 0 \quad (3.12)$$

Entretanto, esse sistema possui infinitas soluções, dado que qualquer coluna da matriz \mathbf{Q} pode ser escrita como o negativo da soma das demais colunas.

Porém, existe uma restrição a mais do sistema que está ligada à seguinte propriedade:

$$\vec{\pi} \cdot \vec{1} = 1 \quad (3.13)$$

em que $\vec{1}$ é um vetor coluna com o número de linhas igual ao número de estados preenchido com uns. Assim, removendo uma coluna qualquer e adicionando ao sistema de equações essa propriedade como uma nova equação, o sistema se torna determinado com solução única e as probabilidades de estados estáticas podem ser obtidas solucionando o seguinte sistema de equações:

$$\vec{\pi} \begin{bmatrix} -\sum_j q_{1,j} & q_{1,2} & \cdots & q_{1,\sigma-1} & 1 \\ q_{2,1} & -\sum_j q_{2,j} & \cdots & q_{2,\sigma-1} & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ q_{\sigma-1,1} & q_{\sigma-1,2} & \cdots & -\sum_j q_{\sigma-1,j} & 1 \\ q_{\sigma,1} & q_{\sigma,2} & \cdots & q_{\sigma,\sigma-1} & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \end{bmatrix} \quad (3.14)$$

3.4 Limitações do método analítico

A obtenção das métricas de desempenho de forma analítica possui a vantagem de ser uma solução exata, considerando as hipóteses feitas. Teoricamente, para um simulador alcançar a precisão de uma estatística calculada analiticamente, ele deveria fazer seu número de iterações tender a infinito, o que é impossível.

Entretanto, existe uma desvantagem no uso da metodologia analítica no cálculo dessas estatísticas. Segundo (KUMAR *et al.*, 2004a), a cardinalidade de um CTMC que descreve o comportamento de um enlace multiclasse cresce com o número de classes seguindo a tendência exponencial mostrada na Figura 19. O mesmo pode ser mostrado para o aumento do número de slots totais do modelo do sistema. Dessa forma, a computação das probabilidades de estado pela Equação 3.14 se torna impraticável quando número de classes ou slots totais se torna excessivamente grande, inviabilizando a utilização do método analítico para a obtenção dos valores desejados, para os tipos de algoritmos que não geram CTMC reversível.

Esse problema pode crescer ainda mais em complexidade se a escolha da rede não for o enlace único. Cada novo enlace é um multiplicador no número total de slots da rede, de forma que uma analogia semelhante à feita anteriormente também pode ser obtida.

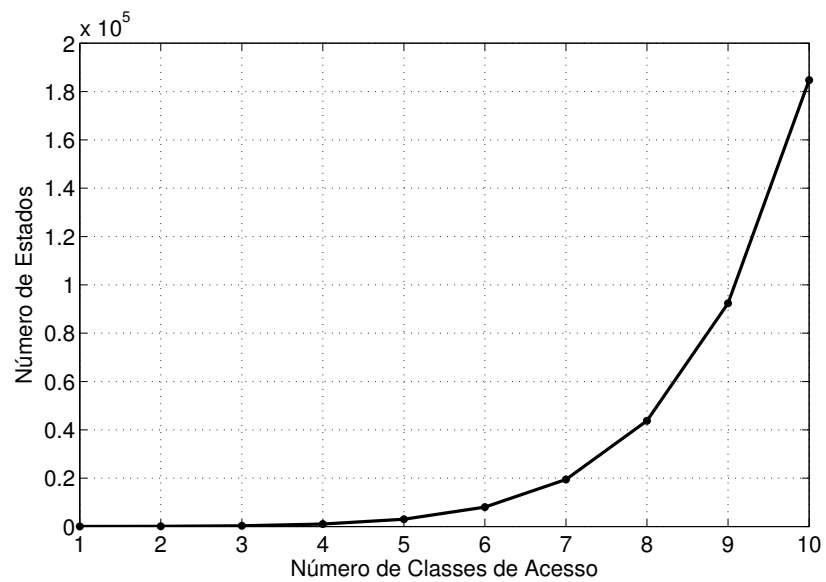


Figura 19 – Explosão do número de estados numa CTMC de acordo com o número de classes de acesso (todas as classes ocupam um slot e o número de slots totais do sistema fixo em dez).

4 Custo da Imparcialidade

Para fazer uma análise quantitativa da perda de eficiência na alocação espectral quando utilizados algoritmos imparciais, foi necessária a criação de uma ferramenta computacional capaz de calcular as métricas de desempenho para diversos casos de estudo.

O estudo de diferentes cenários foi necessário para se criar uma caracterização do comportamento dos algoritmos quando acontecerem mudanças no perfil de tráfego ou para diferentes tamanhos do espectro (escalabilidade).

O capítulo se organiza da seguinte forma: primeiramente, na Seção 4.1, é explicada a ferramenta computacional criada para gerar os resultados; depois, na Seção 4.2, as configurações utilizadas para gerar os resultados são explicitadas; em seguida, são mostrados os resultados obtidos na Seção 4.3 e são feitas análises e comparações dos mesmos.

4.1 Ferramenta computacional

Para se fazer uma análise dos algoritmos de alocação do espectro em termos do desempenho ou imparcialidade de cada um, deve-se, primeiramente, criar um conjunto de métricas que permitam que cada algoritmo seja comparado quantitativamente, determinando, assim, qual é o melhor segundo a métrica escolhida.

Dessa forma, foi criada uma ferramenta computacional matemática capaz de gerar valores a partir de equações, que serão usados posteriormente como métricas para determinar a diferença entre os algoritmos, em termos de tráfego de dados e outros valores.

4.1.1 Funcionamento da ferramenta

A ferramenta é um programa de computador capaz de, dado um par de algoritmos de alocação do espectro, um parcial e o outro imparcial, gerar um conjunto de gráficos das diversas métricas estudadas em função da carga de tráfego normalizada. Esses gráficos comparam os resultados dos dois algoritmos, mostrando o impacto do aumento do tráfego na rede em seus desempenhos.

A normalização da carga de tráfego é necessária para gerar gráficos comparáveis entre os diversos tamanhos de espectro explorados. Para todos os resultados, o tempo de serviço foi mantido fixo em 1 unidade temporal enquanto as taxas de chegadas foram sendo variadas para simular diferentes padrões de tráfego, de baixa carga até alta carga, em erlang por slots. As

equações 4.1 e 4.2 foram utilizadas para obtenção dos valores de carga de tráfego.

$$C_N = \frac{E}{S} \quad (4.1)$$

$$E = \frac{\lambda}{\mu} \quad (4.2)$$

Em que C_N é carga de tráfego normalizada, E é a carga de tráfego de slots em erlang, S é o tamanho do espectro em número de slots de 12,5 GHz, λ é o valor de taxa de chegada de requisições em requisições por unidade de tempo e $1/\mu$ é o tempo médio de vida de uma conexão em unidade temporal.

A ferramenta computacional foi construída utilizando *scripts* para o software MathWorks Matlab e não usa *toolboxes* de terceiros. A ferramenta não tem todas as suas funções em um só *script*, no lugar disso, ela funciona como um conjunto de funções separadas que exercem partes essenciais do processo de criação dos valores utilizados para gerar os gráficos. A Figura 20 traz um fluxograma do funcionamento da ferramenta, mostrando o fluxo da informação que é gerada no decorrer do processo.

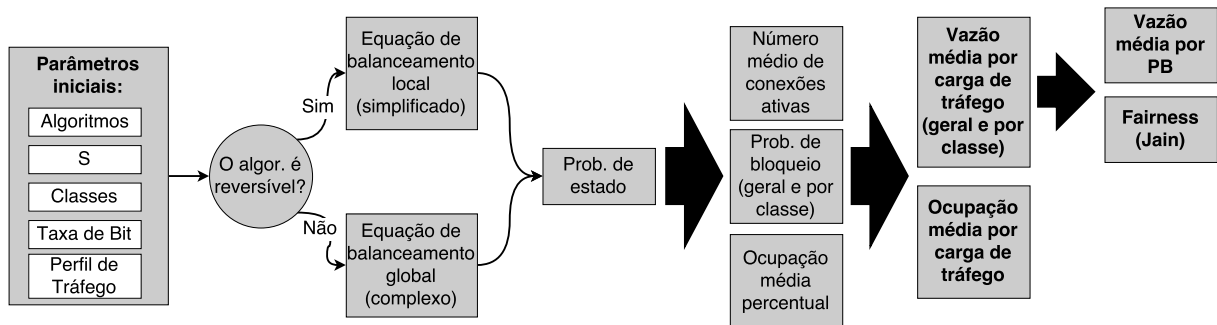


Figura 20 – Fluxograma do funcionamento da ferramenta computacional.

As entradas do programa são: os algoritmos de alocação espectral (um parcial e outro imparcial), o tamanho do espectro (S), as classes de acesso, os valores de vazão por conexão e carga de tráfego (taxa de chegada de requisição e tempo médio de serviço). Os valores pagos pelos usuários de cada classes são considerados proporcionais aos tamanhos de cada classe, por exemplo: se uma conexão da classe A ocupa o dobro de espectro que uma conexão da classe B, um usuário da classe A pagará o dobro que um usuário da classe B.

O módulo inicial do programa decide como calcular as probabilidades de estado. Caso o comportamento do algoritmo de alocação em questão possa ser descrito com uma cadeia de Markov reversível, as probabilidades de estado serão calculadas utilizando a equação de balanceamento local para reduzir o custo computacional segundo as equações apresentadas em (KUMAR *et al.*, 2004a). Caso não possa ser descrito por uma cadeia reversível, as probabilidades deverão ser calculadas utilizando equação de balanceamento global.

Em seguida, a partir dos valores de probabilidade de estado, diversos outros valores podem ser derivados dependendo do resultado que se deseja alcançar. Para cada estado descrito no vetor de probabilidade de estados, o programa armazena um vetor descrevendo o número de conexões de cada classe para aquele estado. Essas informações são utilizadas posteriormente para o cálculo de diversas métricas de desempenho, como mostrado a seguir.

4.1.2 Cálculo das métricas de desempenho

A primeira métrica calculada pelo programa é o número médio de conexões ativas por classe. Para calcular esse valor, deve-se fazer uma média ponderada do número de conexões ativas, em que os pesos são as probabilidades de estado. A equação descrita em seguida descreve o cálculo desse valor:

$$N_m(c) = \sum_{e \in E} N(e, c) \times Pr(e) \quad (4.3)$$

Em que, $N_m(c)$ é o número médio de conexões ativas da classe c , E é o conjunto de todos os estados possíveis, $N(e, c)$ é uma função que retorna o número de conexões ativas da classe c no estado e e $Pr(e)$ é a probabilidade de o sistema estar no estado e . Embora tenha sido dito que essa é uma operação de média ponderada, não existe denominador. Isso se dá pelo fato de o fator ponderador ser a probabilidade de estado, que, quando somada para todos os estados possíveis, resulta em um.

Em seguida, o programa gera os valores de vazão média. Para obtê-los, deve se usar as informações de vazão por conexão para cada classe de tráfego, que são entradas do programa, juntamente com os valores de número médio de conexões ativas, calculados anteriormente. O cálculo desse valor é descrito pelas equações a seguir:

$$V(c) = N_m(c) \times V_{conex}(c) \quad (4.4)$$

$$V_m = \sum_{c \in C} V(c) \quad (4.5)$$

Em que, $V(c)$ representa o valor de vazão média em GB/s da classe c , $V_{conex}(c)$ é o valor de vazão por conexão da classe c e V_m descreve o valor de vazão média total.

O terceiro valor computado é a probabilidade de bloqueio (PB). Para cada configuração de entrada, o programa computa todos os estados possíveis. Os estados que não aceitam novas requisições de uma dada classe, são marcados como estados bloqueantes para aquela classe. Assim, o valor da probabilidade de bloqueio de uma classe, que descreve qual a probabilidade de o sistema recusar uma nova conexão dessa classe, é a soma das probabilidades dos estados bloqueantes para a classe em questão.

Dessa forma, para calcular a PB de uma dada classe, foi utilizada a equação a seguir:

$$PB(c) = \sum_{e \in E} Pr(e) \times Bloq(e, c) \quad (4.6)$$

Em que $PB(c)$ é a PB para a classe c e $Bloq(e, c)$ é uma função booleana que retorna 1 para estados que bloqueiam a classe c e 0 para estados que não a bloqueiam.

A partir desses valores, o programa calcula os valores de PB média utilizando a seguinte equação:

$$PB_m = \frac{\sum_{c \in C} PB(c) \times TaxaReq(c) \times TaxaBit(c)}{\sum_{c \in C} TaxaReq(c) \times TaxaBit(c)} \quad (4.7)$$

Em que, $TaxaReq(c)$ é uma função que retorna a taxa de chegada de requisição de conexões da classe c em requisições por unidade temporal e $TaxaBit(c)$ é uma função que retorna a taxa de bit de uma conexão da classe c em GB/s.

O quarto valor que o programa gera é o de ocupação média do espectro. Ele é inicialmente calculado para cada classe e, então, é calculado o valor da ocupação total. Esse conjunto de valores é dado em fração do espectro e é obtido segundo as equações a seguir:

$$Ocup(c) = \sum_{e \in E} Pr(e) \times \left(\sum_{c \in C} \frac{tam(c) \times N(c, e)}{S} \right) \quad (4.8)$$

Em que $Ocup(c)$ simboliza a ocupação média da classe c em fração do tamanho do espectro e $tam(c)$ é uma função que retorna o tamanho, em número de slots, de uma conexão da classe c . Assim, a ocupação total do espectro é dada por:

$$Ocup_t = \sum_{c \in C} Ocup(c) \quad (4.9)$$

Em que $Ocup_t$ descreve a ocupação total de todas as classe em fração do espectro. Para o particionamento, como o tráfego dentro de cada partição pode ser considerado separadamente como um fila finita, o valor ocupação de cada classe pode ser obtido com uma modificação da Lei de Little (LITTLE; GRAVES, 2008):

$$Ocup(c) = \frac{N_m(c) \times tam(c)}{S} = \frac{\lambda_c \times tam(c)}{\mu_c \times S} \quad (4.10)$$

Em que λ_c e μ_c são, respectivamente, a taxa de chegada de requisições da classe c e a taxa de término de conexões da classe c .

Por último, a ferramenta gera o valor de índice de imparcialidade segundo o índice de Jain (Equação 2.1). Para tanto, utiliza-se dos valores de vazão como recurso a ser compartilhado

e as relações entre as taxas de bit como custo, considerando o valor pago por cada cliente é proporcional à taxa de bit que ele recebe. Assim, as equações a seguir descrevem o cálculo do índice de imparcialidade:

$$Imp = \frac{[\sum_{c \in C} a(c)]^2}{n_{classes} \times \sum_{c \in C} a(c)^2} \quad (4.11)$$

$$a(c) = \frac{V(c)}{TaxaBit(c)/\min(TaxaBit(C))} \quad (4.12)$$

Em que Imp é o valor de índice da imparcialidade calculado, $n_{classes}$ é o número de classes escolhido nas configurações iniciais, $V(c)$ é a vazão observada pela classe c , $\min(TaxaBit(C))$ é uma função que o retorna a menor taxa de bit dentre as classes de acesso.

A consideração feita é apenas uma das possíveis para preço pago por cada tipo de cliente. Considerando que o dono da infraestrutura queira cobrar por porção do espectro oferecido para cada cliente, o valor pago seria, então, proporcional ao tamanho, em slots, de cada classe. Assim, cliente com maior ocupação pagariam mais que cliente com menor ocupação.

Além disso, a concessionária tem um custo fixo administrativo para cada cliente. Isso deve adicionar um ganho (bias) em cada valor pago, tornando os preços não proporcionais, como discutido anteriormente.

Para a geração das curvas de cada gráfico, a fim de se obter pares de métrica de desempenho por carga de tráfego, o processo descrito no fluxograma é repetido diversas vezes, cada uma para um valor diferente de carga de tráfego.

4.1.3 Obtenção dos Resultados Gráficos

A partir dos valores de vazão média, probabilidade de bloqueio, ocupação média do espectro e índice de imparcialidade, o programa itera dentro do conjunto de valores de carga de tráfego para gerar os seguintes gráficos: vazão versus carga de tráfego, vazão versus probabilidade de bloqueio, ocupação versus carga de tráfego e índice de imparcialidade versus carga de tráfego.

A fim de se comparar o desempenho dos dois algoritmos de alocação imparcial estudados nesse trabalho, haverá gráficos de comparação do algoritmo imparcial com o algoritmo parcial, que será utilizado aqui como referência de melhor desempenho. Entretanto, para se comparar os algoritmos imparciais entre si, alguns gráficos são feitos com valores dos dois algoritmos de alocação.

O primeiro conjunto de resultados gráficos mostrado nos resultados é o que mostra a vazão média em função da carga de tráfego. Nesse tipo de gráfico, o eixo das abscissas representa a carga de tráfego normalizada e o eixo das ordenadas representa os valores de vazão de tráfego

em GB/s. Cada gráfico do conjunto é gerado para um valor de tamanho do espectro diferente. Por gráfico, existem três curvas e dois pontos. As curvas mostram o comportamento de um algoritmo imparcial, de um algoritmo eficiente e a diferença entre as duas. Os pontos mostram os valores de vazão nos quais cada algoritmo tem probabilidade de bloqueio de aproximadamente 10%. Um exemplo desse gráfico é a Figura 23. A principal informação a ser extraída desse gráfico é o comportamento da eficiência de cada algoritmo quando a carga de tráfego aumenta, ou seja, os usuários da rede passam a requisitar mais conexões ou manter conexões mais duradouras. A curva de diferença traz uma informação quantitativa da perda em eficiência no uso de um algoritmo imparcial.

Para fazer uma comparação entre os dois algoritmos imparciais analisados, esse conjunto de gráficos trás um terceiro gráfico: o gráfico de comparação do custo da imparcialidade. Nesse gráfico, os eixos representam os mesmos valores que os outros. Entretanto, as curvas são as curvas de diferença de desempenho dos entre os algoritmos obtidas anteriormente. Mostrando essas diferenças lado a lado, é possível comparar dois algoritmos imparciais utilizando um algoritmo eficiente, o *upper bound*, como referência. Esse gráfico também explicita para quais valores de carga de tráfego normalizada qual é o algoritmo mais apropriado para ser utilizado.

O segundo conjunto de resultados a ser apresentado posteriormente se refere às probabilidades de bloqueio. Primeiramente, são comparadas as evoluções das probabilidades de bloqueio com o aumento da carga de tráfego. Esses resultados mostram a discriminação dos algoritmos parciais com as classes que ocupam mais espectro. Como os valores de probabilidade de bloqueio média por carga de tráfego e de vazão por carga de tráfego são gerados a partir do mesmo conjunto de valores de carga de tráfego (entrada do programa), a ferramenta realiza uma interpolação de primeiro grau para se obter os valores de vazão média relacionados a valores não previamente calculados de probabilidade de bloqueio. As Figuras 21 e 22 explicita em imagens o processo de interpolação.

Assim, em cada gráfico desse tipo existem três curvas: uma para um algoritmo parcial, uma para um algoritmo imparcial e uma para a diferença entre as duas primeiras. Nesse gráfico, a principal informação, assim como no gráfico anterior, é o desempenho de cada algoritmo. Como os algoritmos levam carga de tráfego diferente para alcançar determinados valores de probabilidade de bloqueio, eles podem ser comparados diretamente para cada valor de probabilidade de bloqueio.

Assim, como nos gráficos de vazão por carga de tráfego, também é feita aqui uma comparação entre custos dos algoritmos imparciais. As diferenças são comparadas para cada valor de PB analisada.

O terceiro conjunto de gráficos da seção de resultados é o de ocupação por carga de tráfego. Nele, o eixo das abcissas representa os valores de carga de tráfego normalizada e o eixo das ordenadas demonstra os valores de ocupação média do espectro em porcentagem. Cada gráfico tem uma curva para a ocupação geral do algoritmo, e outras n curvas para cada

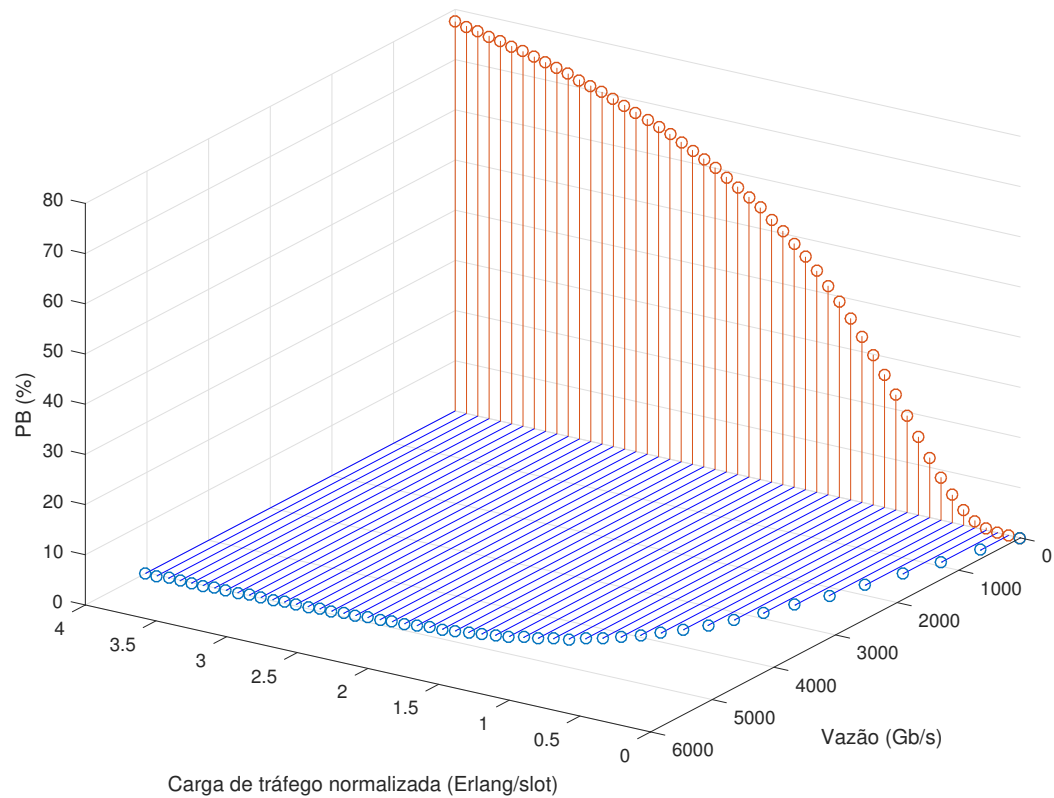


Figura 21 – Relação entre as amostras de vazão média e de PB.

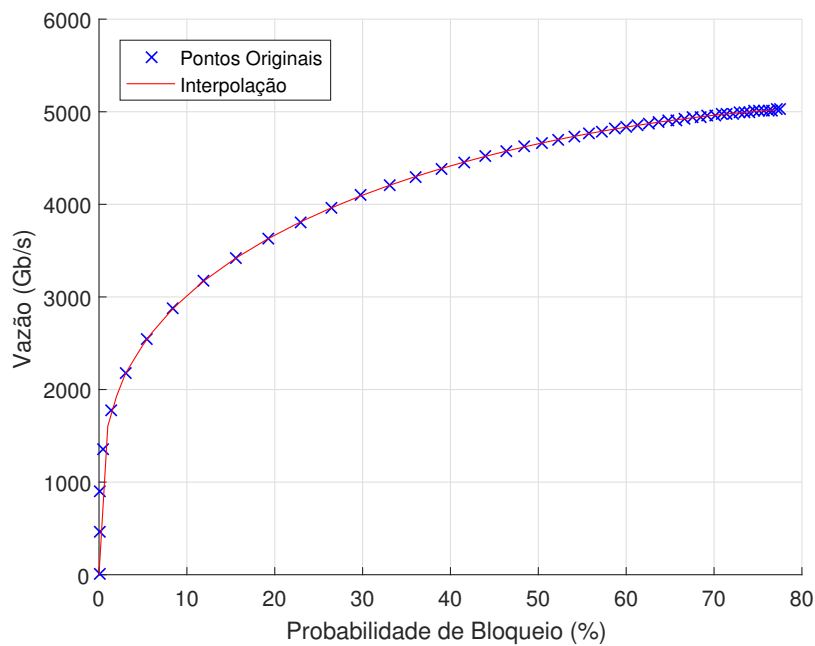


Figura 22 – Interpolação dos valores de Vazão para os valores de PB não presentes nos resultados originais.

classe de acesso. Esse gráfico pode ser visto, por exemplo, na Figura 28. Esse resultado pode ser usado para comparar o quão imparcial é a alocação de cada algoritmo, comparando-se o comportamento de cada classe de acesso com o aumento da carga de tráfego.

Dentre os resultados, o quarto conjunto de gráficos apresentados é o de índice de imparcialidade de Jain por carga de tráfego. Esses gráficos tem duas curvas e dois pontos. Cada par curva-ponto representa um algoritmo de alocação espectral (novamente, um eficiente e um imparcial). Uma amostra pode ser visto na Figura 29. Esses gráficos demonstram, em concordância com os gráficos de ocupação, a imparcialidade na alocação do espectro para alguns tipos de algoritmo e configurações.

4.2 Arranjo Experimental

De forma a tornar os resultados descritos nesse documentos reproduzíveis, essa seção traz uma descrição detalhada dos parâmetros e métodos utilizados em todas as etapas do processo de obtenção dos resultados.

A configuração escolhida foi de enlace único de fibra óptica com extensão de 8000 *km*. A escolha do enlace único, como explicado em (WALDMAN *et al.*, 2015), permite a análise do desempenho dos algoritmos de alocação sem a ação da fragmentação que ocorreria entre os enlaces da rede.

Conforme explicado na Seção 4.1, a metodologia utilizada para a obtenção das probabilidades de estado é analítica. Entretanto, para algumas configurações testadas, o tempo computacional foi excessivamente alto. Esses tempos computacionais mais altos podem ser explicados pela falta de escalonabilidade do método analítico. Como o processamento se baseia em cálculos matriciais, quando essas matrizes aumentam significativamente de tamanho, em decorrência da escolha das configurações, o tempo de processamento também aumenta. Isso limitou as análises para enlaces de até 164 slots. A Figura 19, da Seção 3.4 mostra o crescimento do número de estados de uma CTMC com o aumento do número de classes de acesso, mantido fixo o número total de slots.

O processo de chegada de requisições de conexão foi modelado como estocástico e segue uma distribuição exponencial entre os tempos de chegada de requisições, sendo, assim, denominado um processo de Poisson. Considerando-se que existem K classes de acesso, cada classe tem seu tempo médio entre chegadas definido como λ_k^{-1} , em que k é o índice da k -ésima classe de acesso, ou seja, $k \in \{1, 2, \dots, K\}$, e λ_k é a taxa de chegada. O processo de serviço (tempo de vida de uma conexão) também foi modelado como de Poisson e os tempos médios de conexão das classes k são definidos como μ_k^{-1} . Salvo menção do contrário, será adotado o valor fixo de $\mu_k = 1$ para todo k .

Utilizando o conjunto de resultados apresentados em (SOUZA *et al.*, 2016), foram cria-

das duas situações de investigação. A primeira é um enlace de 8000 Km com classes de acesso de 13 e 28 slots que correspondem a taxas de bit de 400 GB/s e 1 TB/s, respectivamente, e número de slots (S) de 164 slots de 12,5 GHz. O segundo conjunto de resultados utiliza as mesmas configurações de classes de acesso, porém o tamanho do enlace foi reduzido para 163 slots. A tabela 2 faz um resumo das configurações descritas anteriormente. O tamanho de cada slot espectral padronizado pelo ITU-T é de 12,5 GHz, o que gera um número total de slots de aproximadamente 320, dado que a faixa de frequências convencional, banda C, tem aproximadamente 4 THz de banda. Não foi possível gerar resultados para 320 slots devido ao elevado tempo de computação.

Tabela 2 – Configurações dos conjuntos de resultados da Seção 4.3

Conjunto	Tamanho do Enlace	Classes	Taxas de bit (GB/s)	Número de slots
1	8000 Km	13,28	400,1000	164
2	8000 Km	13,28	400,1000	163

Quanto ao perfil de tráfego, as duas classes de acesso tem mesma taxa de chegada e tempo de serviço. Essa escolha permite que seja trivial a configuração do algoritmo imparcial de particionamento do espectro. Para que haja imparcialidade, basta que as partições do espectro destinem o mesmo número médio de conexões para cada classe proporcional ao valor pago por cada categoria de cliente.

Outra escolha no perfil de tráfego foi a escolha dos valores de taxa de chegada, lembrando que os valores de tempo de serviço foram mantidos fixos. Diferentes das outras configurações, que tem somente um valor por resultado, as taxas de chegada são uma lista ordenada crescente de valores. O objetivo é criar um conjunto de resultados para uso posterior na criação de gráficos. Os valores crescentes descrevem o impacto do aumento do uso da rede, simulando momentos de tráfego intenso. Para os resultados descritos na próxima seção, os valores de taxa de chegada foram variados de forma a criar valores de taxa de tráfego normalizada entre 0 e 4 erlang por slot.

4.3 Resultados e análises

4.3.1 Conjunto de resultados 1 (164 slots)

O valor de 164 slots foi escolhido porque, nele, cabem exatamente 4 conexões de cada classe de acesso. Isso facilita a configuração das partições no algoritmo de particionamento. Assim, para esse algoritmo, o espectro fica dividido em duas partições: 52 slots contínuos para a classe de conexões de 13 slots e 112 slots contínuos para as conexões da classe de 28 slots.

Nas Figuras 23.a e 23.b, estão descritos os comportamentos da vazão de tráfego para cada algoritmo em comparação com o algoritmo eficiente. Nesses gráficos, assim como em muitos outros nessa seção, o algoritmo eficiente está representado como Desfragmentação nas

legendas. Apesar de a desfragmentação não ser um algoritmo de alocação, essa legenda indica que ela foi utilizada sobre qualquer algoritmo voraz (*greedy*) de alocação. O algoritmo imparcial de *Trunk Reservation* também usa desfragmentação para que seja criada uma comparação justa entre os algoritmos. No caso do particionamento, a desfragmentação não é necessária dado que, dentro da mesma partição não há perda de fragmentação.

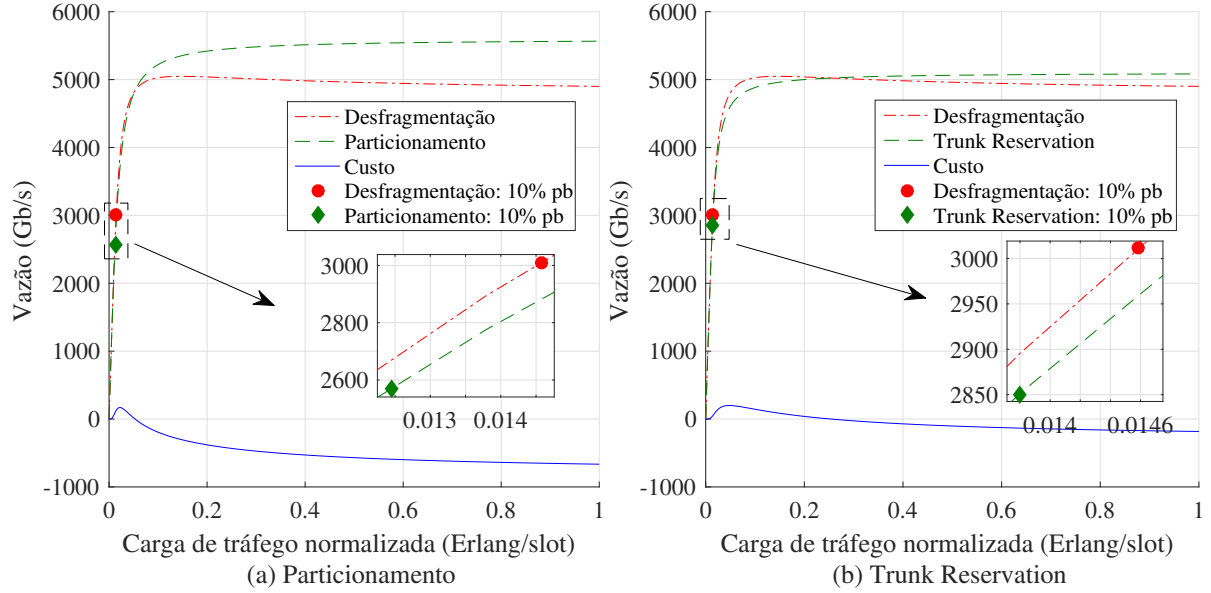


Figura 23 – Vazão por carga de tráfego para $S = 164$ slots.

A partir desses gráficos, foi observado que os algoritmos imparciais, em geral, tem desempenho inferior aos algoritmos eficientes para valores de carga de tráfego baixa. Entretanto, vê-se que, para valores de carga de tráfego alta, a vazão alcançada pelos algoritmos é superior à do algoritmo *greedy*. Isso acontece porque os algoritmos imparciais, para altas cargas de tráfego, aloca uma quantidade igual de conexões de cada classe de acesso. A razão entre o tamanho em número de slots de uma conexão de cada classe é dado por:

$$R_c = \frac{c_2}{c_1} = \frac{28}{13} \approx 2,15... \quad (4.13)$$

Enquanto isso, a razão entre as taxas de bit é dada por:

$$R_t = \frac{t_2}{t_1} = \frac{1000}{400} = 2,5 \quad (4.14)$$

Isso implica que, quando o algoritmo imparcial aloca duas conexões da classe de 1 TB/s (classe 2) em detrimento de cinco conexões da classe de 400 GB/s (classe 1), ele aumenta a vazão por slot porque a mesma vazão de 2 TB/s usa apenas 56 slots no lugar de 65.

Como destacado no *zoom* nos gráficos, para 10% de probabilidade de bloqueio, as vazões alcançadas pelos algoritmos imparciais são menores que pelo algoritmo eficiente.

Como pode ser visto nas curvas de custo da imparcialidade para ambos os algoritmos imparciais, essas assumem valores negativos. Esse “custo negativo” indica um ganho em vazão na escolha dos algoritmos imparciais. Embora isso possa indicar, a princípio, que os algoritmos imparciais são melhores que os algoritmos vorazes, é importante notar que existe uma dependência com o valor de carga de tráfego de slots para existir ganho na escolha do algoritmo imparcial. Esse aumento da carga de tráfego vem com o aumento da probabilidade de bloqueio. Resultados mais detalhados sobre essas relações serão descritos a seguir.

A partir das curvas de custo dos gráficos na Figura 23, a Figura 24 compara os custos da imparcialidade entre os dois algoritmos para as configurações deste conjunto de resultados. Diferentemente do gráfico anterior, os valores de carga de tráfego só variam entre 0 e 0,1 erlang/slot pois é onde se encontra a informação mais pertinente à análise.

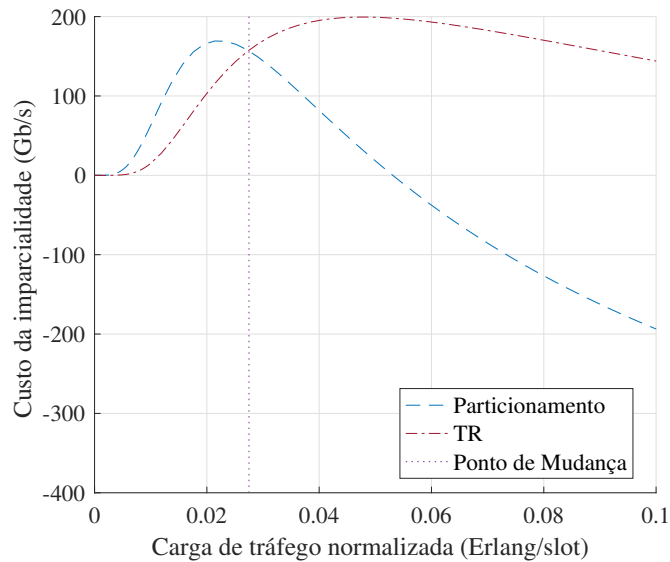


Figura 24 – Comparação do custo da imparcialidade em vazão dos dois algoritmos imparciais para $S = 164$ slots.

Algoritmos com maior custo tem desempenho inferior, ou seja, quanto menor o valor de custo, melhor é o algoritmo para a carga analisada. Esse gráfico traz um importante resultado: não existe um algoritmo imparcial que sempre supera o outro em desempenho. Ao invés disso, para carga de tráfego normalizada superior a aproximadamente 0,0275 erlang/slot, ocorre a troca do algoritmo de maior custo. Para valores de carga inferiores a 0,0275 erlang/slot, o algoritmo de *Trunk Reservation* tem desempenho melhor que o algoritmo de Particionamento mas isso se inverte para carga maiores. Esse resultado pode ser explicado pelo fato de que o algoritmo de TR nunca preenche totalmente o espectro, gerando uma perda cada vez maior para carga de tráfego crescente.

Analisando cada curva de custo do tráfego separadamente, é possível ver que ambos algoritmos têm um comportamento crescente inicialmente até um ponto de máximo com queda em seguida. Isso pode ser explicado pelo fato de que, ao tentar dar uma porcentagem do espec-

tro para cada classe proporcional à sua ocupação, o tráfego final é maior do que somente alocar conexões eficientemente, pelos motivos explicados anteriormente. Essa inversão de comportamento só se manifesta para valores de carga de tráfego normalizada de aproximadamente 0,0214 erlang/slot para o algoritmo de Particionamento e 0,0461 erlang/slot para o algoritmo de TR. Os custos máximos observados foram de 169,3 GB/s e 199,3 GB/s para os algoritmos de Particionamento e de TR, respectivamente.

O próximo conjunto de resultados se refere às probabilidades de bloqueio. O gráfico da Figura 25 mostra as probabilidades de bloqueio de cada classe para cada algoritmo de alocação analisado.

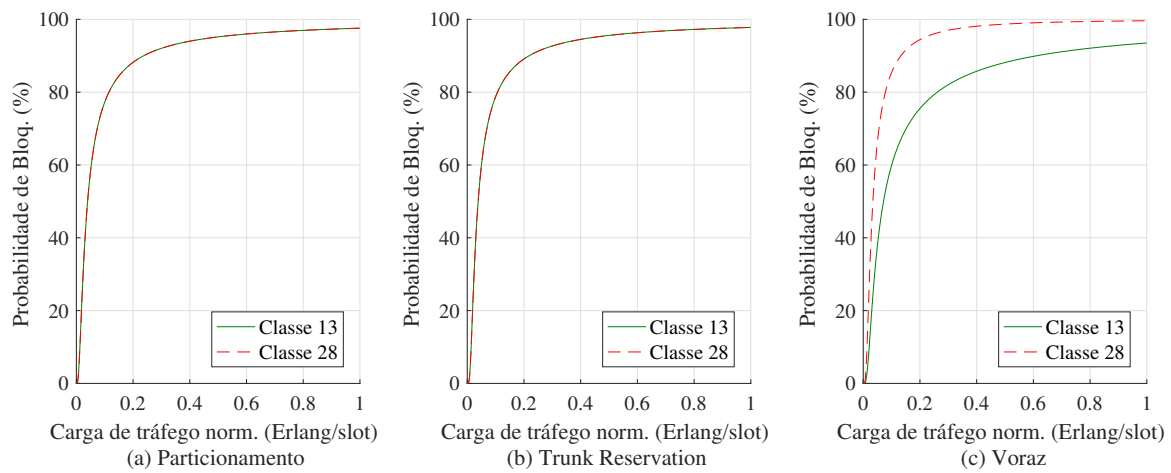


Figura 25 – Probabilidade de bloqueio dos algoritmos de alocação analisados para $S = 164$ slots.

O comportamento dessa métrica para o algoritmo eficiente é mostrado na Figura 25.c. Pode-se notar que as curvas de cada classe estão sempre sempre separadas, com a PB da classe de 28 slots sempre maior. Esse comportamento mostra que a alocação eficiente bloqueia mais as conexões de maior ocupação, classe de 28 slots no gráfico.

Em contraste, vê-se nas Figuras 25.a e 25.b que os algoritmos imparciais geram probabilidades de bloqueios iguais para as classes de acesso, que pode ser interpretado como uma medida de geração de imparcialidade.

Com os valores de PB e de vazão dos gráficos anteriores, nas Figuras 26.a e 26.b, é possível ver o comportamento dos dois algoritmos em relação ao aumento da probabilidade de bloqueio.

Como as probabilidades de bloqueio aumentam com a carga de tráfego (Figura 25), os resultados desses gráficos concordam com os comportamentos apresentados para vazão por carga de tráfego. O que esse tráfego traz de adicional é a seleção dos maiores valores de vazão alcançáveis para cada algoritmo de alocação. Nota-se, aqui, que o algoritmo eficiente tem um

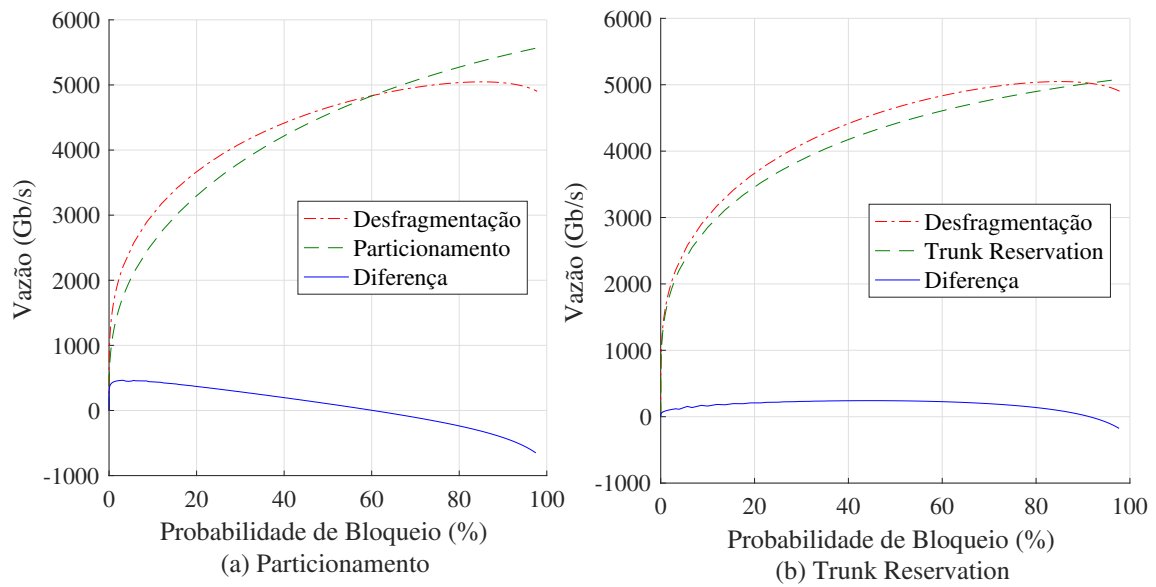


Figura 26 – Vazão por PB para $S = 164$ slots.

distanciamento maior do algoritmo de particionamento para PB de até 50%. Enquanto isso o algoritmo de TR apresenta um comportamento mais semelhante ao visto anteriormente.

Assim como foi feito para os gráficos de vazão por carga de tráfego, a Figura 27 compara as curvas de diferença para os dois algoritmos imparciais. Os valores de vazão negativos representam situações em que o algoritmo imparcial consegue alcançar vazão maior que o algoritmo eficiente.

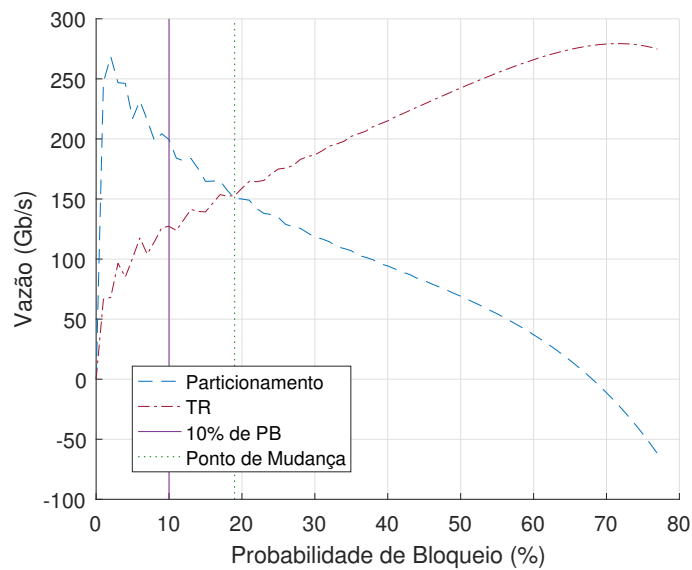


Figura 27 – Comparação do custo da imparcialidade para vazão por PB $S = 164$ slots.

Os resultados mostrados nesse gráfico, apesar das diferenças citadas anteriormente estão

em concordância com o visto no gráfico para vazão por carga de tráfego. O ponto de mudança entre o algoritmo de melhor desempenho ocorre em torno de 35,75% de PB. Para 10% de PB, tem-se o algoritmo de TR com desempenho superior, apresentando um ganho de aproximadamente 279,2 GB/s em relação ao algoritmo de Particionamento.

Outro importante comportamento apresentado pelas curvas de vazão versus PB, é inclinação das curvas quando a probabilidade de bloqueio se aproxima de 100%. Isso indica que, quanto mais próximo do estado de bloqueio total de todas as requisições de conexão, melhor os algoritmos imparciais se comportam, em comparação aos algoritmos vorazes.

O próximo conjunto de gráficos apresentam a ocupação percentual do espectro por carga de tráfego normalizada. Eles estão descritos nas Figuras 28.a (Particionamento), 28.b (*Trunk Reservation*) e 28.c (algoritmo Eficiente). Embora tenham sido gerados resultados para carga de tráfego normalizada entre 0 e 1 erlang/slot, nesses gráficos, só foram mostrados os valores entre 0 e 0,2 pois é a região de comportamento transitório.

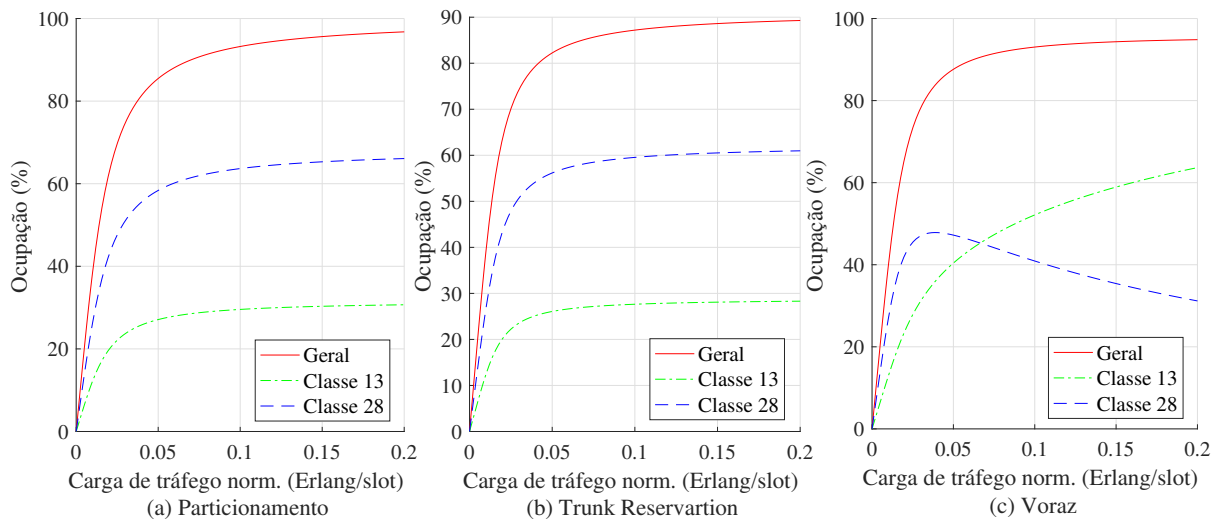


Figura 28 – Ocupação por carga de tráfego para $S = 164$ slots.

As curvas em vermelho nos três gráficos mostram o comportamento da ocupação geral para cada algoritmo. Neste gráfico, quanto maior a ocupação do espectro alcançada para um valor de definido de carga de tráfego, melhor é aquele algoritmo, dado que há menos desperdício no uso do espectro. Enquanto o algoritmo de Particionamento tem um desempenho bastante parecido com o algoritmo eficiente em ocupação média geral, o algoritmo de *Trunk Reservation* apresenta desempenho inferior aos dois, devido à construção de seu algoritmo, sempre deixando de alocar conexões para tornar as probabilidades de bloqueio semelhantes.

Para esse gráfico, como esperado, o algoritmo eficiente supera os outros em quase todas as instâncias. Para a carga máxima analisada o algoritmo de Particionamento ultrapassa o algoritmo eficiente em alocação do espectro. Isso acontece porque, quando a taxa de chegada se

torna muito superior que a taxa de serviço, o algoritmo eficiente tende a alocar somente conexões de tamanho menor (13 slots). Quando uma conexão maior (28 slots) acaba, como as taxas de chegada são iguais (tráfego uniforme), existe uma probabilidade de 50% de chance dessa conexão ser substituída por uma conexão da classe de 13. Desse estado, considerando o regime de tráfego alto e que todos os outros slots estão ocupados, somente uma conexão de 13 slots pode ser alocada, resultando na troca de uma conexão de 28 slots por duas de 13 slots. Assim, o algoritmo eficiente acaba reduzindo a ocupação final do espectro, em relação ao algoritmo de Particionamento, que aloca quantidade proporcional do espectro para cada classe.

Analizando as curvas em azul e verde, ocupações das classes de 28 (1 TB/s) e 13 slots (400 GB/s), respectivamente, foi concluído que os algoritmos imparciais conseguem manter porções proporcionais do espectro para cada classe. Essa proporção segue a razão entre o número de slots de cada classe (Equação 4.13). Por exemplo para carga de tráfego normalizada igual a 0, 2:

$$R_{ocup,partic} = \frac{Ocup_{28}}{Ocup_{13}} = \frac{66,18}{30,73} = 2,15...$$

$$R_{ocup,TR} = \frac{Ocup_{28}}{Ocup_{13}} = \frac{61,03}{28,33} = 2,15...$$

As diferenças nos valores das razões são encontrados somente na terceira casa decimal e são devido a arredondamentos feitos nos cálculos e elas tendem a diminuir quando a carga de tráfego tende ao valor máximo.

Por fim, a Figura 29 mostra o comportamento do índice de imparcialidade de Jain dos algoritmos para a configuração escolhida.

Como esperado, dado que estão operando em situações ideais, ambos os algoritmos imparciais apresentam índice de imparcialidade igual a 1 (perfeita). Esse valor indica que não há discriminação das classes em nenhum momento, mesmo com o aumento da carga de tráfego.

Analizando a área em *zoom*, é possível ver os valores de índice de imparcialidade e carga de tráfego quando as probabilidades de bloqueio são de 10%. Embora apresente um índice inferior a 1 no momento, ou seja, esteja alocando mais espectro de forma desigual, o algoritmo eficiente (descrito com Desfragmentação na figura) apresenta melhor resultado pois alcança uma carga de tráfego maior que os outros dois.

Outro importante resultado extraído do gráfico é o comportamento de algoritmo imparcial. Ele rapidamente decresce em imparcialidade, pois dá prioridade às conexões de baixa ocupação. Entretanto, mostra um comportamento assintótico com aumento da carga de tráfego. Isso indica que, embora haja desigualdade, o algoritmo não tende a alocar a 100% dos recursos para uma classe só quando a carga tende a valores muito altos.

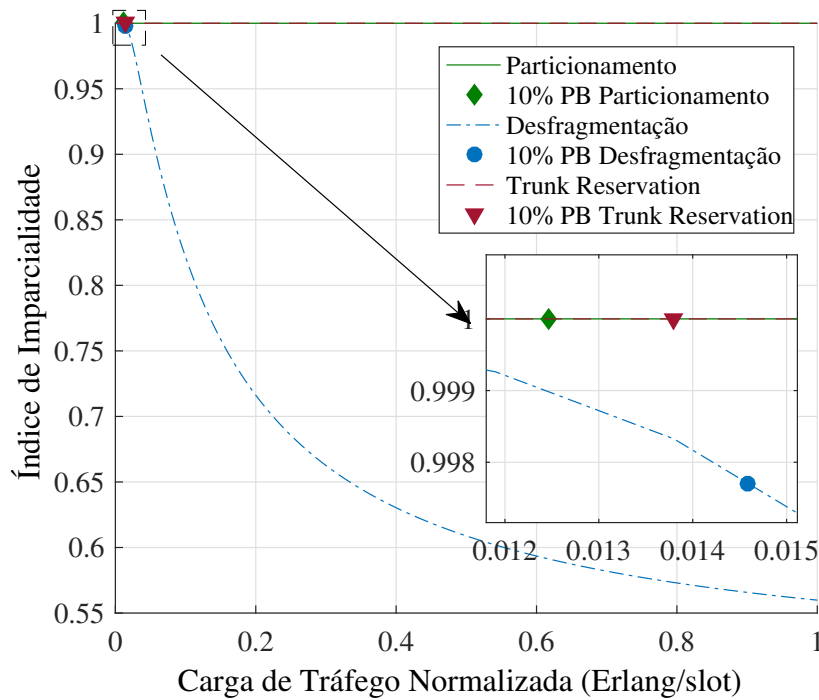


Figura 29 – Imparcialidade por carga de tráfego para $S = 164$ slots.

4.3.2 Conjunto de resultados 2 (163 slots)

O segundo conjunto de resultados foi gerado para uma configuração de espectro de 163 slots espectrais de 12,5 GHz. Esse valor do número de slots foi escolhido para ser feita uma análise do impacto do tamanho no espectro no desempenho dos algoritmos imparciais. Todas as outras configurações foram mantidas iguais às escolhidas para a obtenção dos resultados anteriores.

A Figura 30 mostra os gráficos de vazão por carga de tráfego para essa configuração.

Para o particionamento, o desempenho apresentando para 163 slots foi inferior ao apresentado para a configuração anterior, com custo máximo acima de 1000 GB/s. Entretanto, ainda mantém a propriedade custo decrescente com o aumento da carga de tráfego. A Figura 31 compara o particionamento para os dois tamanhos de espectro. As áreas dos retângulos são proporcionais ao tamanho da partição de cada classe. Para 164 slots, cabem exatamente 4 conexões de cada classe, assim, o algoritmo de particionamento consegue criar as partições que somam o tamanho do espectro. Entretanto, para 163 slots, o algoritmo só pode alocar 3 conexões de cada, de forma que possa se manter imparcial. Assim, somente 123 slots são ocupados por partições, deixando os outros 40 slots (aprox. 24,5% do espectro) não utilizados.

Já para o algoritmo de TR, tem-se que seu desempenho foi inferior ao algoritmo eficiente para todas as cargas de tráfego, porém, vê-se uma tendência decrescente na diferença (custo) entre os dois algoritmos, como observado anteriormente.

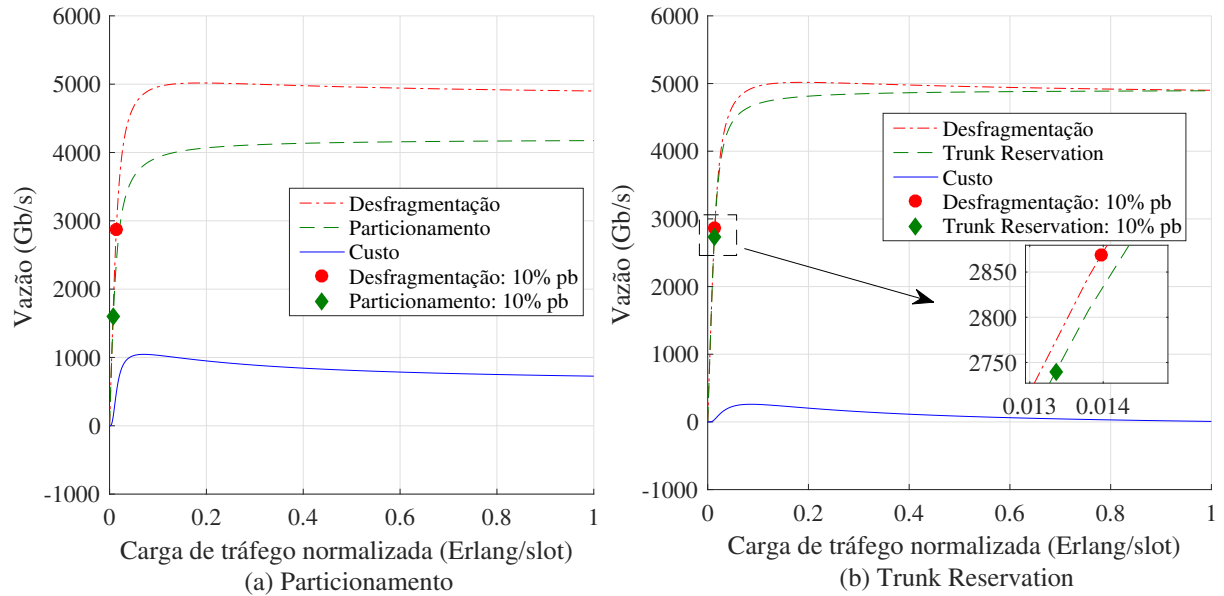


Figura 30 – Vazão por carga de tráfego para $S = 163$ slots.

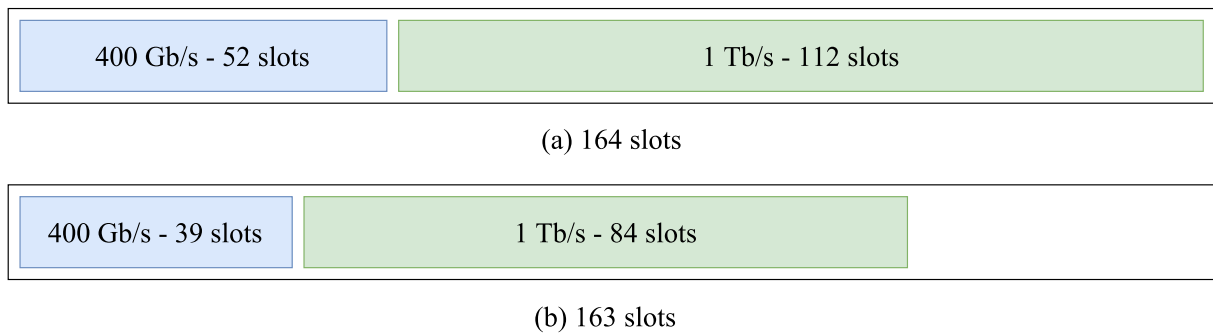


Figura 31 – Comparação do particionamento para os dois tamanhos de espectro analisados.

A Figura 32 mostra a comparação dos custos da imparcialidade para o tamanho de espectro escolhido.

Diferentemente do que visto anteriormente, o algoritmo de Particionamento apresenta desempenho inferior ao TR para todos os valores de carga de tráfego de slots observados. O algoritmo de Particionamento tem custo máximo de 1045 GB/s enquanto que, para o algoritmo de TR, esse valor é de 260,3 GB/s. Esses pontos de custo máximo ocorrem para valores de carga normalizada de 0,07 e 0,08 erlang/slot, para os algoritmos de Particionamento e TR, respectivamente.

Outra grande diferença nesses resultados é que, para o particionamento, o custo máximo é equivalente à perda de 1 conexão de 1 TB/s ou 2 de 400 GB/s.

Em seguida, de forma a caracterizar o comportamento das probabilidades de bloqueios para os algoritmos, tem-se os gráficos da Figura 33.

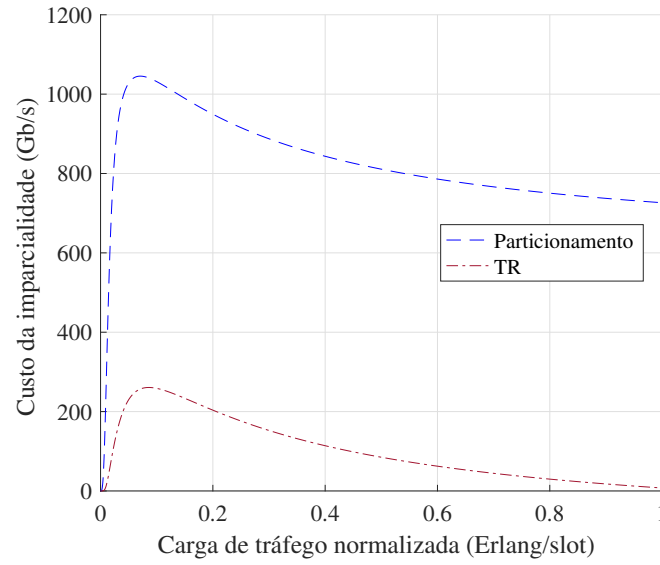


Figura 32 – Comparação do custo da imparcialidade em vazão para $S = 163$ slots.

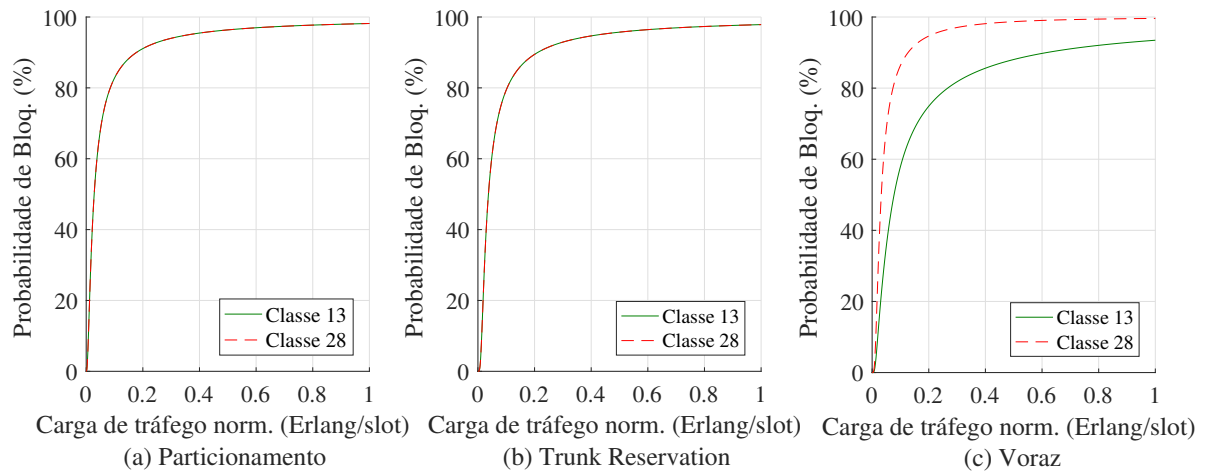


Figura 33 – Probabilidade de bloqueio dos algoritmos de alocação analisados para $S = 163$ slots.

Para esse gráfico, a principal diferença em relação aos resultados anteriores é a perda de desempenho para o particionamento apresentada aqui na forma de uma curva que tende mais rapidamente para 100%.

A partir desses resultados e dos resultados descritos na Figura 30, foram gerados os gráficos na Figura 34, na qual pode-se ver o comportamento da vazão alcançada por cada algoritmo.

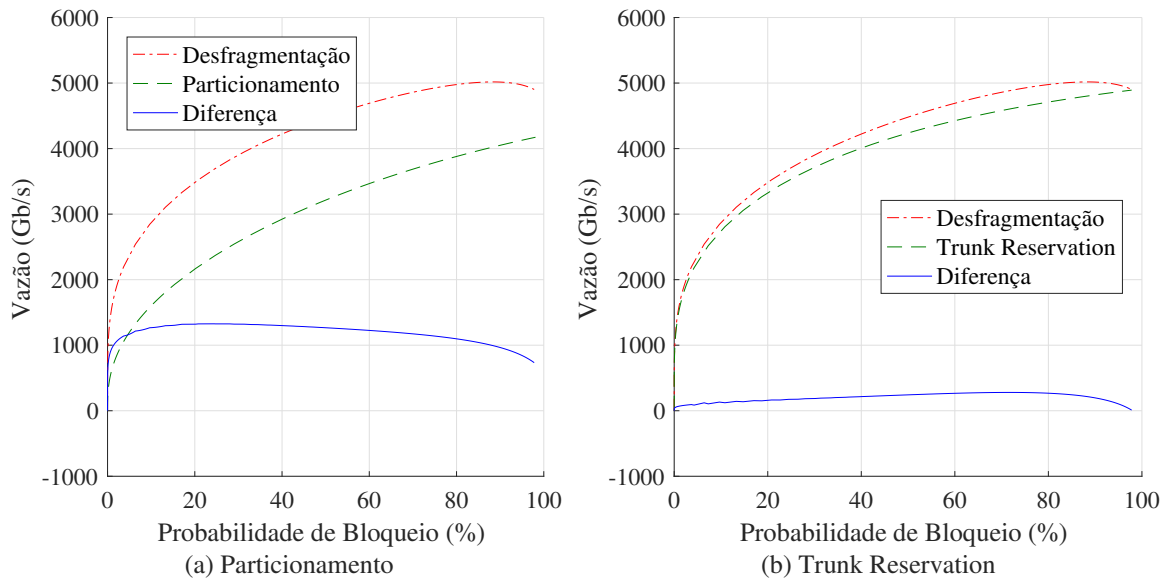


Figura 34 – Comparação do custo da imparcialidade para vazão por PB $S = 163$ slots.

Novamente, o comportamento apresentado pelos algoritmo de Particionamento demonstra perda de desempenho. É interessante notar que nenhum algoritmo alcança custo negativo, ou seja, eles não se tornam melhores que o algoritmo voraz para nenhum valor de probabilidade de bloqueio alcançado pelos resultados.

A Figura 35 mostra os dois custos da imparcialidade para que possam comparados.

Para 10% de PB, o algoritmo de Particionamento apresenta custo de 1267 GB/s enquanto que o mesmo valor de PB para o algoritmo de *Trunk Reservation* tem custo de 128 GB/s. Isso pode ser entendido como um ganho de desempenho na escolha do último algoritmo de 1139 GB/s, dado que se deseja trabalhar num regime de 10% de PB. Esse resultado mostra o quanto o algoritmo de Particionamento sofre com o tamanho total do espectro escolhido, em detrimento do TR que quase não sofre alteração no desempenho.

Para se fazer uma análise inicial da imparcialidade na alocação de cada algoritmo, a Figura 36 traz as ocupações percentuais do espectro que cada algoritmo alcança.

Aqui, novamente, a queda de performance do algoritmo de Particionamento fica evidente. Diferentemente de anteriormente, o particionamento apresenta a menor ocupação, fi-

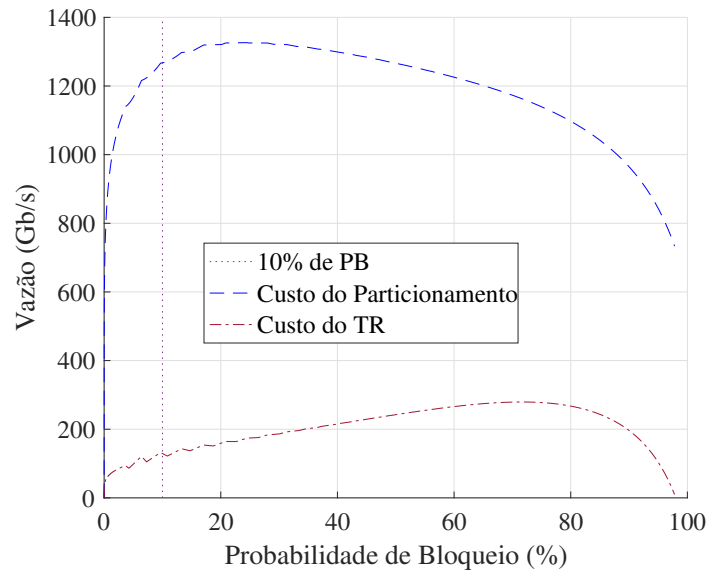


Figura 35 – Comparação do custo da imparcialidade para vazão por PB $S = 163$ slots.

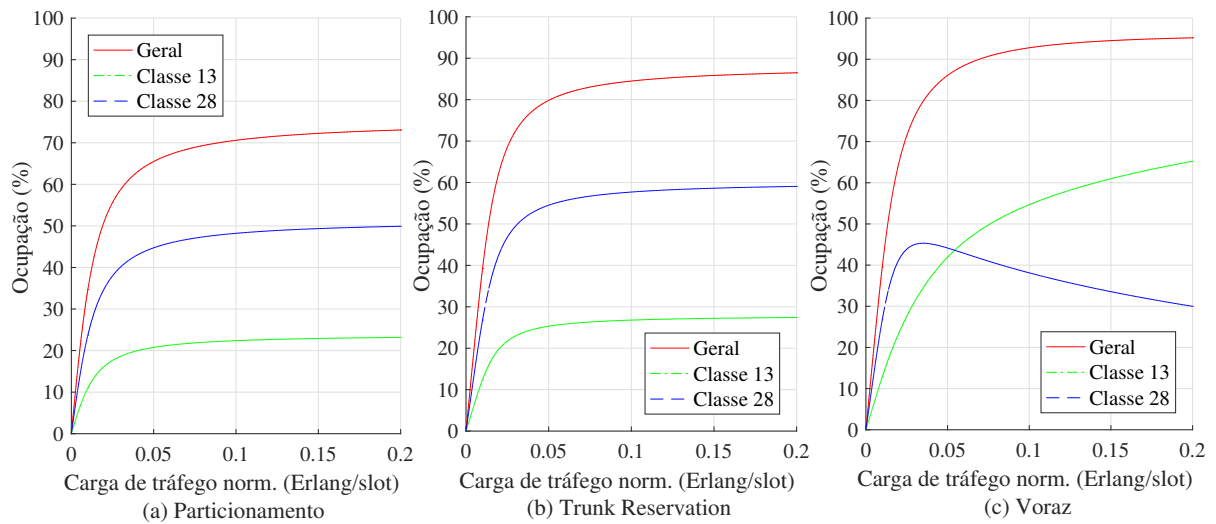


Figura 36 – Ocupação por carga de tráfego para $S = 163$ slots.

cando com ocupações inferiores ao *Trunk Reservation* e ao algoritmo voraz. Apesar disso, o algoritmo ainda apresenta ocupação proporcional entre as duas classes.

Por fim, a Figura 37 traz a representação gráfica do comportamento do índice de imparcialidade de Jain calculado de acordo com a Equação 4.11 para os três algoritmos.

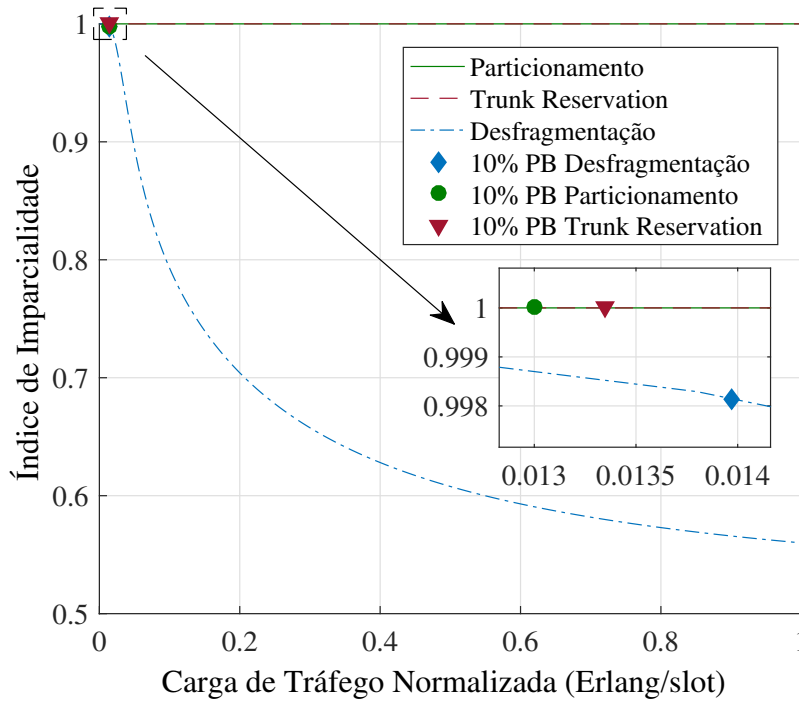


Figura 37 – Imparcialidade por carga de tráfego para $S = 163$ slots.

Nesse resultado, fica evidente que ambos os algoritmos conseguem manter a imparcialidade para qualquer configuração de espectro. Além disso, o algoritmo de particionamento encontra 10% de PB muito antes que os outros dois algoritmos analisados. Esse ultimo resultado é explicado pelo fato de, ao somente alocar conexões em uma área inferior a 100% do espectro, o Particionamento permanece mais tempo em estados bloqueantes.

4.3.3 Conjunto de resultados 3 (Pseudo-particionamento)

Como visto na seção anterior, o algoritmo de Particionamento do espectro sofre graves perdas de desempenho de acordo com a relação entre o número máximo de conexões do espectro e os tamanhos das conexões das classes de acesso.

Esse problema pode ser amenizado substituindo-se o algoritmo de Particionamento pelo algoritmo de Pseudo-particionamento (Seção 2.3.1). Isso porque o segundo algoritmo tenta utilizar 100% do espectro para alocação de conexões, o que o torna mais eficiente.

Para mostrar esse ganho de desempenho, a Figura 38 traz a vazão média obtida pelo algoritmo de Pseudo-particionamento versus a carga de tráfego para o espectro com 163 slots.

Quando comparada com a Figura 30.a, vê-se que a primeira agrega mais vazão por carga chegando até a ultrapassar o algoritmo voraz.

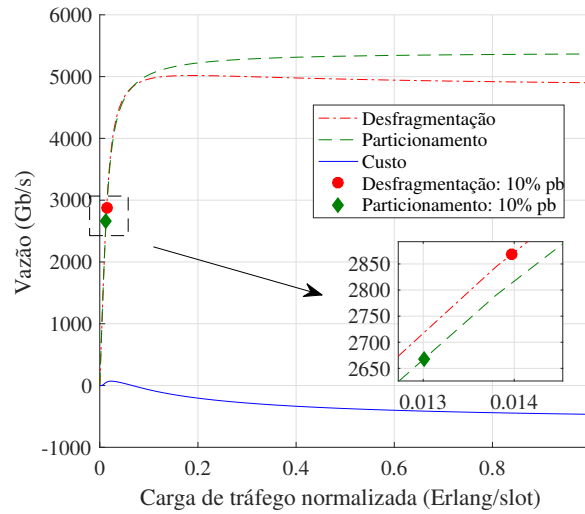


Figura 38 – Vazão versus carga de tráfego para $S = 163$ slots (Pseudo-particionamento)

Esse ganho pode ser explicado analisando a Figura 39. Ela compara graficamente a alocação de conexões do Particionamento com a alocação do Pseudo-particionamento.

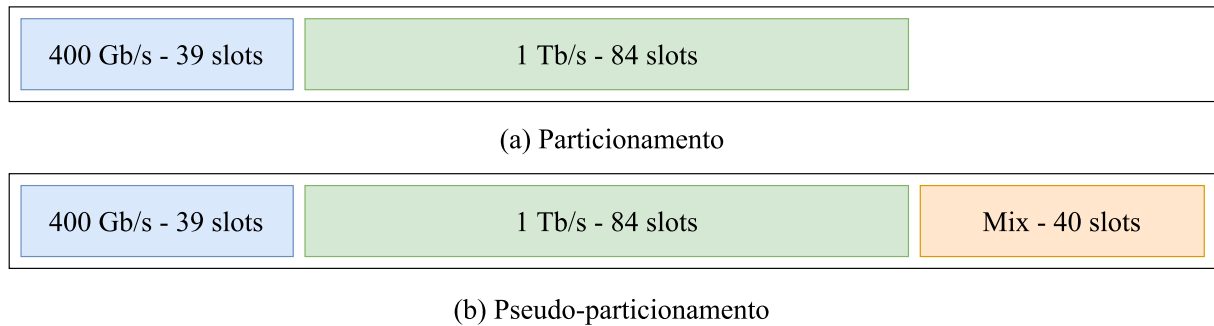


Figura 39 – Comparação das partições entre o Particionamento e o Pseudo-particionamento para $S = 163$ slots.

De forma semelhante ao algoritmo de particionamento, o Pseudo-particionamento também cria duas partições de uso exclusivo de cada classe de acesso, garantindo uma porção proporcional do espectro para cada uma. Porém, ele cria uma terceira partição, representada pelo termo “Mix” na figura. Essa partição, que nesse caso tem 40 slots, aceita conexões das duas classes de acesso, o que permite agregar o tráfego excedente das outras partições, ao invés de bloqueá-lo. Entretanto, essa partição deve ser desfragmentada para evitar perdas de fragmentação.

A Figura 40 mostra o comportamento do índice de imparcialidade de Jain para valores crescentes de carga de tráfego.

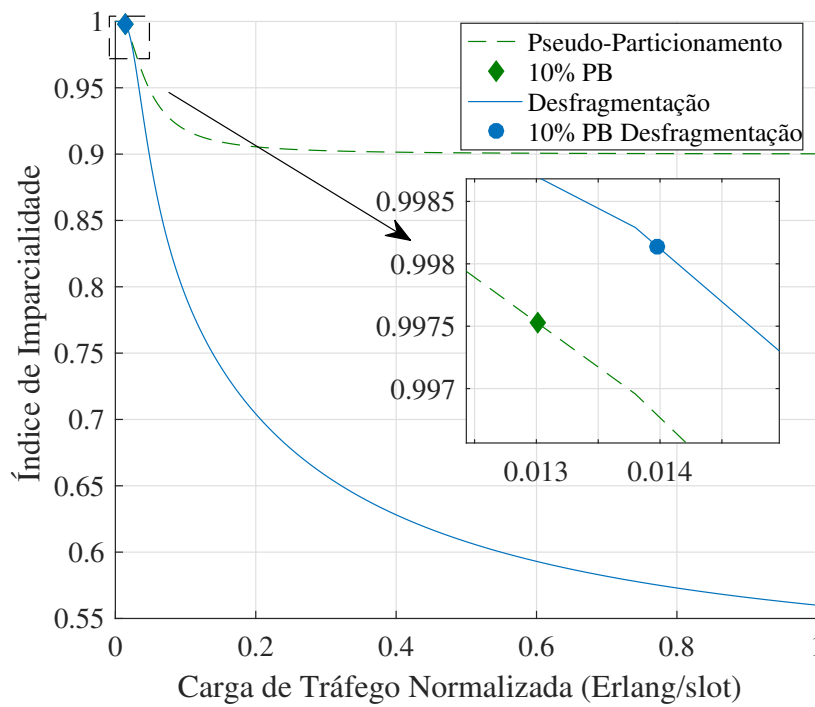


Figura 40 – Índice de imparcialidade de Jain para $S = 163$ slots (Pseudo-particionamento)

Como pode ser observado, o valor do índice inicia em 1, para valores baixos de carga de tráfego, e depois apresenta comportamento assintótico para o valor de 0,9. Isso demonstra a falta de imparcialidade do algoritmo, que tenderá a alocar mais conexões de 13 slots na partição adicional do que conexões de 28 slots. Esse resultado é diferente do apresentado pelo algoritmo de Particionamento, que sempre apresenta imparcialidade perfeita.

Entretanto, em comparação com o algoritmo voraz (Figura 37), o Pseudo-particionamento apresenta um comportamento mais imparcial. Isso pode ser explicado pela criação das duas partições de acesso exclusivo de cada classe, que cria uma distribuição mais igualitária do espectro do que o livre acesso por todas as classes.

5 Comportamento assintótico

5.1 Efeito de aumento do tráfego nas cadeias de Markov

No capítulo anterior, foi investigado o comportamento de diversas métricas em relação ao aumento gradual da carga de tráfego. Essas métricas foram obtidas até a carga normalizada de 1 erlang por slot. Isso porque, a partir desse valor, os bloqueios são inevitáveis devido à falta de slots para acomodar tal demanda, mesmo se o sistema funcionasse sob regime determinístico.

Entretanto, esses valores de carga não estudados anteriormente são plausíveis e podem acontecer em momentos de alta demanda. Um gerente de rede pode estar interessado em projetar seu sistema para não sofrer perdas excessivas nessas ocasiões. Por isso, é interessante investigar o que acontece com o aumento gradativo da carga para valores além de 1 erlang por slot.

A partir dos resultados do capítulo anterior, nota-se que o comportamento da curva do custo da imparcialidade em tráfego de dados é composta de duas regiões para qualquer dos algoritmos de alocação imparcial testados e até para o pseudo-particionamento. Em uma primeira região, para valores baixos de carga de tráfego, ocorre um crescimento até um valor de pico do custo. Isso ocorre porque esses algoritmos bloqueiam conexões sem necessariamente estarem com o espectro totalmente ocupado.

Todavia, com o aumento da carga de tráfego, a lógica desses algoritmos de alocação passa a favorecer o tráfego acumulado, dado que as conexões de maior ocupação geralmente possuem maior eficiência espectral. Assim, depois do pico de custo, esse valor começa a diminuir.

Para ilustração desse comportamento, essa seção não usará as configurações dos resultados da seção anterior, dado que as cadeias de Markov dos casos estudados no capítulo anterior possuem muitos estados. Dessa forma, como feito no Capítulo 3, será tomado como referência um sistema com duas classes de acesso, uma de 2 slots por conexão e uma de 4 slots por conexão.

As Figuras 41, 42 e 43 mostram, para a mesma configuração de \vec{b} e C , as cadeias de Markov equivalentes ao funcionamento de cada algoritmo estudado. Esses grafos são semelhantes aos mostrados anteriormente no Capítulo 3, diferindo somente pelo destaque para a cadeia de Markov assintótica do sistema quando a carga de tráfego aumenta tendendo a infinito, mantida a relação entre λ_1 e λ_2 . Essa nova cadeia é obtida pela eliminação dos estados com transição de saída rotulada por λ_1 e/ou por λ_2 , uma vez que o tempo de permanência do sistema nesses estados tenderá então a zero, fazendo com que a probabilidade desses estados se aproxime assintoticamente de zero. Restarão então apenas os estados que bloqueiam todas as requisições ((0, 3), (2, 2), (4, 1) e (6, 0)), uma vez que os tempos de vida das conexões ativas

se mantêm inalterados. Supondo $\lambda_1 = \lambda_2$, ao transitar do estado $(0, 3)$ para $(0, 2)$ com taxa 3, o sistema retorna instantaneamente para $(0, 3)$ com probabilidade $1/2$, e com probabilidade $1/2$ segue instantaneamente para $(1, 2)$ e dali para $(2, 2)$, onde permanece por tempo médio positivo, gerando uma transição assintoticamente equivalente de $(0, 3)$ para $(2, 2)$ com taxa $3/2$, como mostrado na Figura 41. A mesma dinâmica gera as transições assintóticas de $(2, 2)$ para $(4, 1)$ com taxa 1, e de $(4, 1)$ para $(6, 0)$ com taxa $1/2$. Porém, quando o sistema transita de $(6, 0)$ para $(5, 0)$, ele imediatamente volta para $(6, 0)$ e lá permanece por um tempo positivo, caracterizando o estado $(6, 0)$ como assintoticamente absorvente.

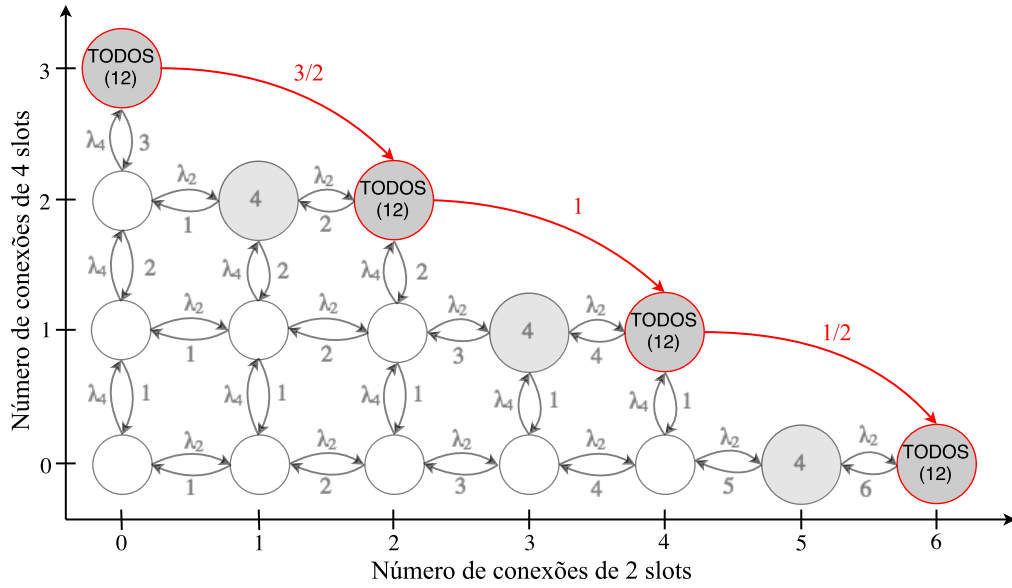


Figura 41 – Cadeia de Markov para altas carga de tráfego utilizando um algoritmo voraz de alocação para $\vec{b} = [2, 4]^T$, $C = 12$ slots.

Após a remoção dos estados transitórios (estados não circunscritos em vermelho nas Figuras 41, 42 e 43), ficarão somente aqueles que só possuem saídas com taxas finitas. Entretanto, para algumas situações como a do algoritmo voraz (Figura 41), alguns estados não apresentam realimentação, de forma que, uma vez que sair desses estados, o sistema virtualmente não retornará a eles.

Assim, é possível concluir que, para esse exemplo, o sistema ficará a maior parte do tempo no estado de 12 conexões de 2 slots para o algoritmo voraz, no estado de 2 conexões de 2 slots e 2 conexões de 4 slots no particionamento e transitará entre diversos estados para o caso do TR.

Analisando os gráficos de ocupação e vazão do Capítulo 4, observa-se nas curvas de custo, um comportamento assintótico, dado que, com o aumento do tráfego, os sistemas irão transitar menos entre todos os estados possíveis, tendendo a passar a maior parte do tempo em estados de ocupação máxima. Pode ser notado que os estados de ocupação máxima nos algorit-

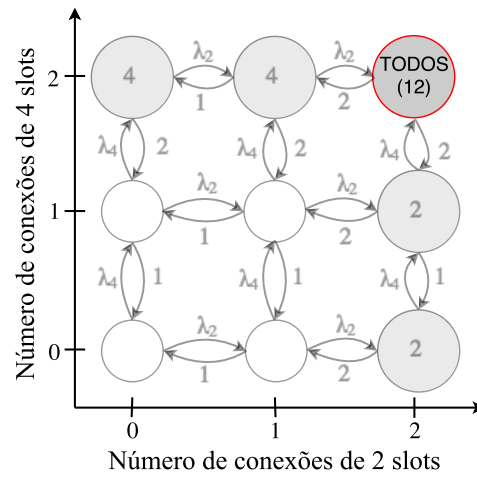


Figura 42 – Cadeia de Markov para altas carga de tráfego utilizando um algoritmo de particionamento para $\vec{b} = [2, 4]^T$, $C = 12$ slots.

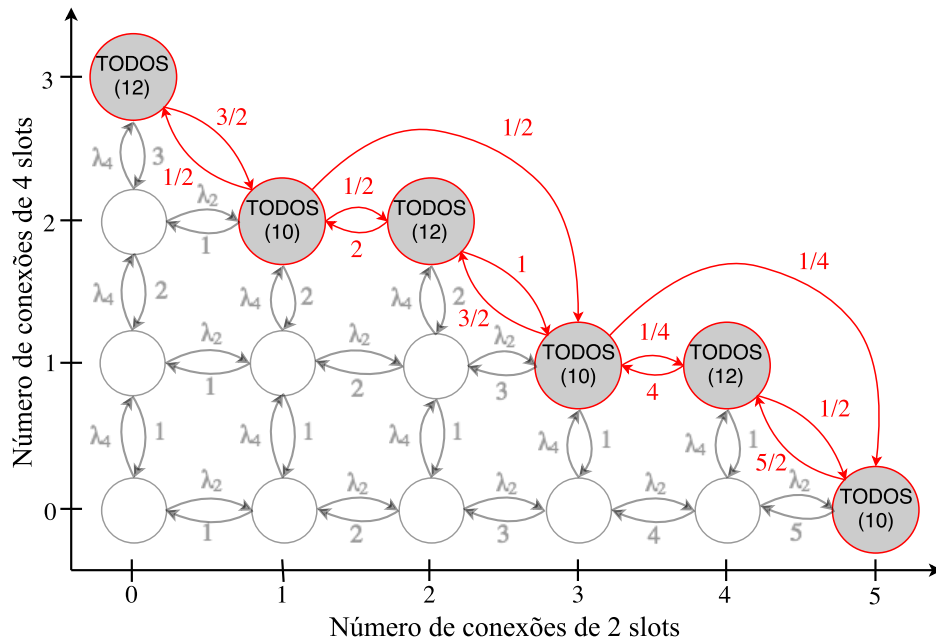


Figura 43 – Cadeia de Markov para altas carga de tráfego utilizando um algoritmo TR para $\vec{b} = [2, 4]^T$, $C = 12$ slots.

mos vorazes e no particionamento apresentam uma ocupação média maior que a do algoritmo TR. Dessa forma, os valores de custo de ocupação do algoritmo TR tendem a um valor maior quando o tráfego cresce a valores muito altos.

A obtenção do valor numérico da assíntota pode ser feito resolvendo o sistema de equações provenientes das CTMCs como os destacadas em vermelho nas Figuras 41, 42 e 43. A próxima seção mostra como o valor da assíntota é alterado pelo tamanho do espectro.

5.2 Maximização da vazão

Na seção anterior, a análise do comportamento assintótico considerou somente a ocupação do espectro pelas classes de acesso. Continuando a análise de vazão de informação, iniciada nos resultados do Capítulo 4, essa seção irá contextualizar as análises de comportamento assintótico com os valores de ocupação espectral e taxa de bit, contidas em (SOUZA *et al.*, 2016), para gerar informações de vazão de tráfego.

Dessa forma, no caso de estudo abordado, as classes de acesso formam um vetor $\vec{b} = [13, 28]^T$ e serão estudados os casos em que $C = 164$ slots e $C = 163$ slots. Nas Figuras 44, 45 e 46 estão demonstradas as “cadeias de Markov assintóticas” para $C = 164$ slots e, quando possível, as alterações na CTMC que ocorrem quando $C = 163$ slots.

Devido à complexidade da CTMC gerada, a Figura 46 apresenta somente o caso de $C = 164$ slots. O caso de $C = 163$ slots pode ser obtido pela remoção do estado de 52 slots ocupados pela classe de ocupação 13 e 112 slots ocupados pela classe de ocupação 28, que é o único que soma 164 slots de ocupação total.

Como os algoritmos vorazes (Figura 44) e de particionamento (Figura 45) tendem a ficar somente em um estado, a análise do valor assintótico do custo do algoritmo de particionamento é uma simples diferença entre a vazão desses dois estados.

Entretanto, como a CTMC resultante para o TR (Figura 46) não é composta de somente um estado, seu valor de custo não é trivial. Utilizando a técnica de obtenção das probabilidades de estado estacionárias para cadeias não-reversíveis (Capítulo 3) nas CTMC resultantes, é possível obter os valores assintóticos descritos nas Figuras 47 e 48. Essas Figuras mostram as curvas de custo da imparcialidade por carga de tráfego normalizada observadas no Capítulo 4 e os valores de assíntota obtidos.

Devido à natureza discretizada das possíveis configurações de número de slots e as CTMCs resultantes, os algoritmos alternam entre qual apresenta desempenho superior. Entretanto, nota-se que o algoritmo de TR transiciona de forma mais suave, apresentando desempenho semelhante para ambos os casos.

Essa propriedade pode justificar o uso do algoritmo de TR ao invés do particionamento, dado que seu desempenho pode ser superior e ele sempre mantém imparcialidade perfeita, sem

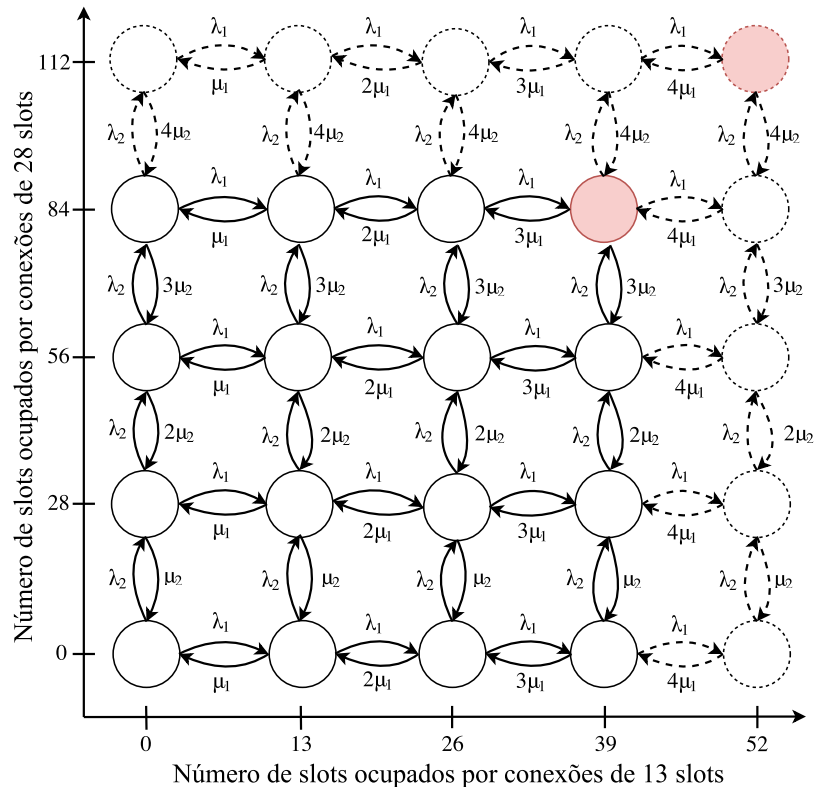


Figura 45 – Cadeia de Markov para altas carga de tráfego utilizando um algoritmo de particionamento para $\vec{b} = [13, 28]^T$, $C = 164$ slots e $C = 163$ slots. Os estados e transições em tracejados são exclusivos do caso em que $C = 164$ slots.

desperdiçar muito espectro nem exigir uma lógica complexa de implementação.

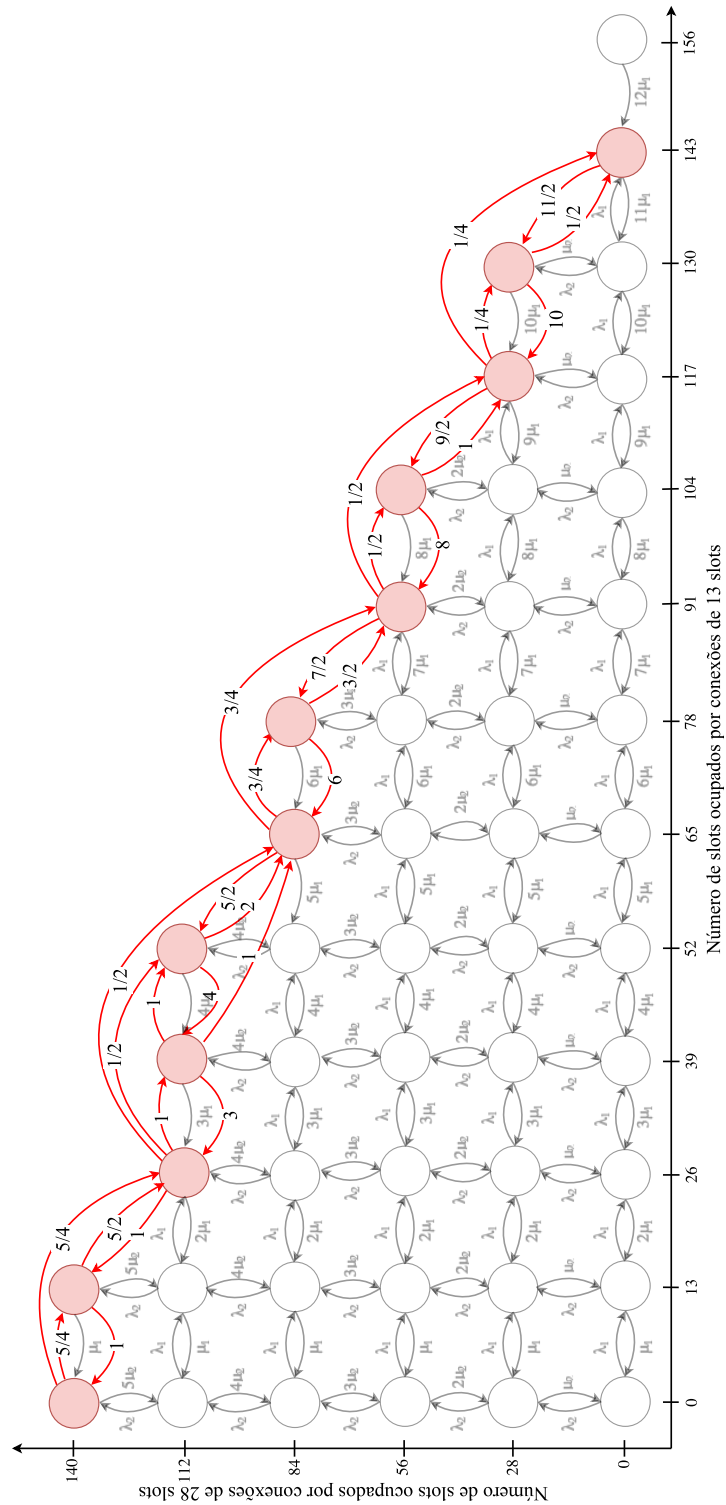


Figura 46 – Cadeia de Markov para altas carga de tráfego utilizando um algoritmo TR para $\vec{b} = [13, 28]^T$, $C = 164$ slots.

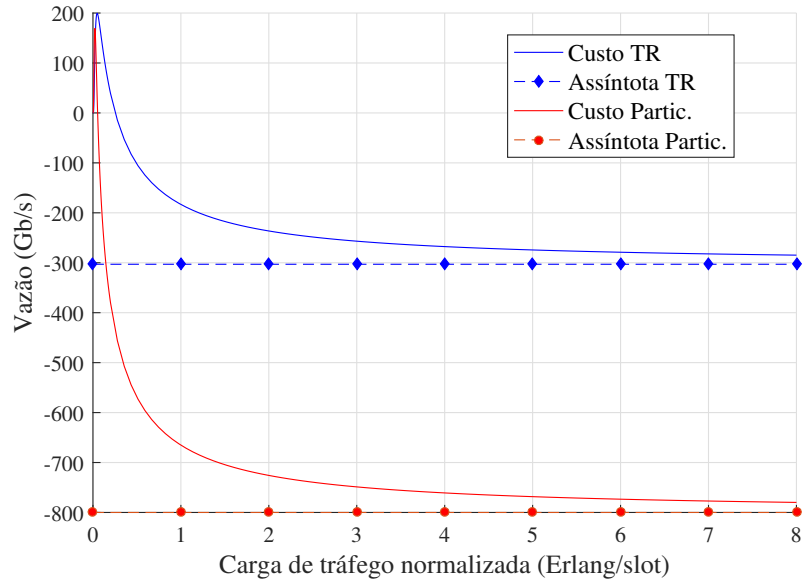


Figura 47 – Custo da imparcialidade assintótico para $\vec{b} = [13, 28]^T$ e $S = 164$ slots.

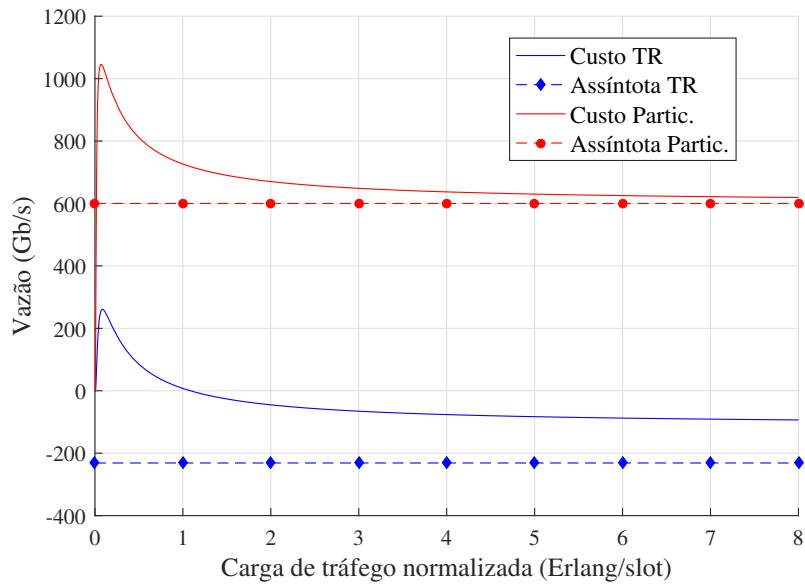


Figura 48 – Custo da imparcialidade assintótico para $\vec{b} = [13, 28]^T$ e $S = 163$ slots.

Conclusões

Esse trabalho comparou o desempenho de dois algoritmos de alocação de espectro imparciais, quantificando o trade-off entre imparcialidade e eficiência. O caso investigado foi o enlace único para que os efeitos da fragmentação entre enlaces não afetasse a performance dos algoritmos. Também foi considerado padrão de tráfego uniforme entre as classes, de forma que a imparcialidade perfeita fosse alcançada para o algoritmo de particionamento, permitindo que ele “empatasse” com o TR neste quesito. Por fim, foi considerada a aplicação de técnica de desfragmentação para simplificar a obtenção dos resultados.

Os resultados foram obtidos analiticamente, usando como referência valores de vazão e ocupação espectral empíricos, utilizando como referência os valores otimizados de configuração apresentados em (SOUZA *et al.*, 2016), foram escolhidas três configurações de forma que pudesse ser estudado o impacto da proporção entre a banda total do sistema e o espectro ocupado pelas classes de acesso. Esses resultados apontam para o fato de que certas configurações de taxa de bit e ocupação de cada classe de acesso podem se beneficiar do uso de algoritmos imparciais ao invés de algoritmos vorazes, aumentando a ocupação e/ou a vazão global do sistema.

As comparações entre os dois tipos de algoritmo imparcial mostram que o algoritmo de particionamento apresenta melhor desempenho em configurações que possibilitem as partições cobrirem 100% do espectro. Isso foi evidenciado para os casos em que o número de slots do sistema é um múltiplo da soma das ocupações de cada classe. Para os outros casos, o algoritmo de TR pode apresentar desempenho semelhante ou superior, além de ser de implementação mais simples e versátil. Isso foi evidenciado pela pequena mudança de desempenho entre as configurações testadas.

Por fim, foi visto que o aumento do número de slots (ou a diminuição nos tamanhos de conexão por classe) gera uma diminuição progressiva no custo de usar ambos os algoritmos imparciais. Além disso, foi visto que os algoritmos transicionam entre qual apresenta melhor desempenho, sendo o TR o algoritmo de desempenho mais estável. Isso foi demonstrado com a análise do comportamento do custo da imparcialidade em vazão de tráfego com a mudança no tamanho total de slots do sistema.

Perspectivas Futuras

O próximo passo na investigação desse tópico, seria investigar as configurações intermediárias entre os valores do tamanho do espectro que geram imparcialidade com aproveitamento total do espectro no particionamento, caracterizando a transição entre a dominância de

um algoritmo ou outro, para um dado vetor \vec{b} (ocupações individuais de cada classe de acesso).

Como essa investigação implica o cálculo de probabilidade de estado em cadeias de Markov de muitos estados, diminui a viabilidade do método analítico, dado que a sua complexidade computacional cresce com o número total de slots do modelo, número de classes de acesso e o tamanho de uma conexão de cada classe de acesso (inversamente). Dessa forma, para obter esses valores no contexto de uma operação real de uma rede óptica elástica de aproximadamente 320 slots de 12,5 GHz, deve ser utilizada simulação de eventos para obtenção das métricas de desempenho.

Referências

- AGRAWAL, G. P. *Fiber-Optic Communication Systems*. Wiley, 2012. (Wiley Series in Microwave and Optical Engineering). ISBN 9780470922828. Disponível em: <<https://books.google.com.br/books?id=yGQ4n1-r2eQC>>. Citado na página 16.
- BERTHOLD, J.; MEMBER, S.; SALEH, A. A. M.; BLAIR, L.; SIMMONS, J. M.; MEMBER, S.; EVOLUTION, A. N. S. Optical Networking : Past , Present , and Future. v. 26, n. 9, p. 1104–1118, 2008. Citado na página 24.
- BERTSIMAS, D.; FARIAS, V. F.; TRICHAKIS, N. The Price of Fairness. *Operations Research*, v. 59, n. 1, p. 17–31, 2011. ISSN 0030-364X. Disponível em: <<http://pubsonline.informs.org/doi/abs/10.1287/opre.1100.0865>>. Citado na página 18.
- BONALD, T.; MASSOULIÉ, L. Impact of fairness on Internet performance. *ACM SIGMETRICS Performance Evaluation Review*, v. 29, n. 1, p. 82–91, 2001. ISSN 01635999. Citado na página 18.
- CALLEGATI, F.; BONANI, L. H.; CERRONI, W. Service Fairness in Flexible Optical Networks. *Optical Fiber Communication Conference/National Fiber Optic Engineers Conference 2013*, n. 1, 2013. Disponível em: <<http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=6532864>>. Citado 2 vezes nas páginas 19 e 30.
- CALLEGATI, F.; BONANI, L. H.; LEZAMA, F.; CERRONI, W.; CAMPI, A.; CASTAÑÓN, G. Trunk reservation for fair utilization in flexible optical networks. *IEEE Communications Letters*, v. 18, n. 5, p. 889–892, 2014. ISSN 10897798. Disponível em: <<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6787150>>. Citado na página 19.
- CHATTERJEE, B. C.; FADINI, W.; OKI, E. A spectrum allocation scheme based on first-last-exact fit policy for elastic optical networks. *Journal of Network and Computer Applications*, v. 68, p. 164–172, 4 2016. ISSN 10958592. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1084804516300017>>. Citado na página 20.
- CHATTERJEE, B. C.; OKI, E. Performance Evaluation of Spectrum Allocation Policies for Elastic Optical Networks. *ICTON 2015*, 2015. ISSN 10897798. Citado na página 20.
- CHATTERJEE, B. C.; SARMA, N.; OKI, E. Routing and Spectrum Allocation in Elastic Optical Networks: A Tutorial. *IEEE Communications Surveys and Tutorials*, v. 17, n. 3, p. 1776–1800, 2015. ISSN 1553877X. Disponível em: <<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=7105364>>. Citado na página 28.
- CUGINI, F.; PAOLUCCI, F.; MELONI, G.; BERRETTINI, G.; SECONDINI, M.; FRESI, F.; SAMBO, N.; POTÍ, L.; CASTOLDI, P. Push-pull defragmentation without traffic disruption in flexible grid optical networks. *Journal of Lightwave Technology*, v. 31, n. 1, p. 125–133, 2013. ISSN 07338724. Citado na página 29.
- DIJKSTRA, E. W. A note on two problems in connexion with graphs. *Numerische Mathematik*, v. 1, n. 1, p. 269–271, 1959. ISSN 0029599X. Citado na página 24.

- FIORANI, M.; TORNATORE, M.; CHEN, J.; WOSINSKA, L.; MUKHERJEE, B. Optical Spatial Division Multiplexing for Ultra-High-Capacity Modular Data Centers. In: *Optical Fiber Communication Conference*. Optical Society of America, 2016. p. Tu2H.2. Disponível em: <<http://www.osapublishing.org/abstract.cfm?URI=OFC-2016-Tu2H.2>>. Citado na página 36.
- GERSTEL, O.; JINNO, M.; LORD, A.; YOO, S. J. B. Elastic optical networking: A new dawn for the optical layer? *IEEE Communications Magazine*, v. 50, n. 2, p. 12–20, 2012. ISSN 01636804. Citado 2 vezes nas páginas 18 e 25.
- JAIN, R.; CHIU, D.; HAWE, W. A Quantitative Measure Of Fairness And Discrimination For Resource Allocation In Shared Computer Systems. 1998. Disponível em: <<http://arxiv.org/abs/cs/9809099>>. Citado 2 vezes nas páginas 30 e 35.
- KADOHATA, A.; HIRANO, A.; FUKUTOKU, M.; OHARA, T.; SONE, Y.; ISHIDA, O. Multi-layer greenfield re-grooming with wavelength defragmentation. *IEEE Communications Letters*, v. 16, n. 4, p. 530–532, 2012. ISSN 10897798. Citado na página 29.
- KELLY, F. P.; MAULLOO, A. K.; TAN, D. K. H. Rate control for communication networks: shadow prices, proportional fairness and stability. *Journal of the Operational Research Society*, v. 49, n. 3, p. 237–252, 1998. ISSN 0160-5682. Disponível em: <<http://link.springer.com/10.1057/palgrave.jors.2600523>>. Citado na página 18.
- KIM, J.; WANG, X.; WANGYAN, S.; RAZO, M.; TACCA, M.; FUMAGALLI, A. Blocking fairness in two-service EONs with uneven arrival rates. *International Conference on Transparent Optical Networks*, v. 2015-Augus, p. 3–6, 2015. ISSN 21627339. Citado na página 20.
- KIM, J.; YAN, S.; FUMAGALLI, A.; OKI, E.; YAMANAKA, N. An analytical model of spectrum fragmentation in a two-service elastic optical link. In: *2015 IEEE Global Communications Conference, GLOBECOM 2015*. [S.l.: s.n.], 2016. ISBN 9781479959525. Citado na página 20.
- KLINKOWSKI, M.; WALKOWIAK, K. On the advantages of elastic optical networks for provisioning of cloud computing traffic. *IEEE Network*, 2013. ISSN 08908044. Citado na página 17.
- KUMAR, A.; MANJUNATH, D.; KURI, J. Circuit-Multiplexed Networks. In: *Communication Networking*. [s.n.], 2004. p. 265–322. ISBN 9780124287518. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/B9780124287518500069>>. Citado 4 vezes nas páginas 39, 43, 52 e 55.
- KUMAR, A.; MANJUNATH, D.; KURI, J. Discrete Event Random Processes. In: *Communication Networking*. [s.n.], 2004. p. 819–865. ISBN 9780124287518. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/B9780124287518500203>>. Citado na página 41.
- LITTLE, J. D. C.; GRAVES, S. C. Chapter 5 Little ’ s Law. *Operations Management*, v. 115, n. November, p. 81–100, 2008. Disponível em: <<http://www.springerlink.com/content/n7h5375373hh8kh1/>>. Citado 2 vezes nas páginas 40 e 57.
- NAG, A.; TORNATORE, M.; MUKHERJEE, B. Optical network design with mixed line rates and multiple modulation formats. *Journal of Lightwave Technology*, v. 28, n. 4, p. 466–475, 2010. ISSN 07338724. Citado na página 16.

- RUAN, L.; ZHENG, Y. Dynamic Survivable Multipath Routing and Spectrum Allocation in OFDM-Based Flexible Optical Networks. *J. Opt. Commun. Netw.*, v. 6, n. 1, p. 77–85, 2014. ISSN 1943-0620. Disponível em: <<http://jocn.osa.org/abstract.cfm?URI=jocn-6-1-77>>. Citado na página 24.
- SARIDIS, G. M.; ALEXANDROPOULOS, D.; ZERVAS, G.; SIMEONIDOU, D. Survey and evaluation of space division multiplexing: From technologies to optical networks. *IEEE Communications Surveys and Tutorials*, v. 17, n. 4, p. 2136–2156, 2015. ISSN 1553877X. Citado na página 35.
- SOUZA, A. L. N.; RUIZ, E. J. M.; REIS, J. D.; CARVALHO, L. H. H.; OLIVEIRA, J. R. F.; ARANTES, D. S.; COSTA, M. H. M.; MELLO, D. A. A. Parameter Selection in Optical Networks With Variable-Code-Rate Superchannels. *Journal of Optical Communications and Networking*, v. 8, n. 7, p. A152, 2016. ISSN 1943-0620. Disponível em: <<https://www.osapublishing.org/abstract.cfm?URI=jocn-8-7-A152>>. Citado 4 vezes nas páginas 25, 61, 80 e 85.
- STRAUB, S.; KIRSTADTER, A.; SCHUPKE, D. A. Multi-period planning of WDM-networks: Comparison of incremental and EoL approaches. *2006 Second IEEE and IFIP International Conference in Central Asia on Internet*, 2006. Citado na página 35.
- TOMKOS, I.; AZODOLMOLKY, S.; SOLE-PARETA, J.; CAREGLIO, D.; PALKOPOULOU, E. A tutorial on the flexible optical networking paradigm: State of the art, trends, and research challenges. *Proceedings of the IEEE*, v. 102, n. 9, p. 1317–1337, 2014. ISSN 00189219. Disponível em: <<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6824237>>. Citado 2 vezes nas páginas 18 e 27.
- VELASCO, L.; MORALES, F.; GIFRE, L.; CASTRO, A.; DIOS, O. González de; RUIZ, M. On-Demand Incremental Capacity Planning in Optical Transport Networks. *Journal of Optical Communications and Networking*, v. 8, n. 1, p. 11, 2016. ISSN 1943-0620. Disponível em: <<https://www.osapublishing.org/abstract.cfm?URI=jocn-8-1-11>>. Citado na página 35.
- WALDMAN, H.; ALMEIDA, R. C.; BORTOLETTO, R. C.; ASSIS, K. D. An analytical defragmentation bound on the performance of the elastic single link with dynamic traffic. In: *International Conference on Transparent Optical Networks*. IEEE, 2015. ISBN 9781467378802. ISSN 21627339. Disponível em: <<http://ieeexplore.ieee.org/document/7193691/>>. Citado 5 vezes nas páginas 19, 26, 27, 45 e 61.
- WALDMAN, H.; SOUZA, I. D. T. D.; ALMEIDA, R. C.; BORTOLETTO, R. C. A comparative discussion of some fairness-generating schemes in elastic networking. *International Conference on Transparent Optical Networks*, 2016. ISSN 21627339. Citado na página 33.
- WANG, R.; MUKHERJEE, B. Spectrum management in heterogeneous bandwidth optical networks. *Optical Switching and Networking*, v. 11, n. PART A, p. 83–91, 2014. ISSN 15734277. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1573427713000799>>. Citado 2 vezes nas páginas 19 e 32.
- WANG, X.; KIM, J.; YAN, S.; RAZO, M.; TACCA, M.; FUMAGALLI, A. Blocking probability and fairness in two-rate Elastic Optical Networks. In: *2014 16th International Conference on Transparent Optical Networks (ICTON)*. IEEE, 2014. p. 1–4. ISBN 978-1-4799-5601-2. ISSN 21627339. Disponível em: <<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6876707>>. Citado na página 19.

YIN, Y.; ZHANG, M.; ZHU, Z.; YOO, S. J. B. Fragmentation-Aware Routing, Modulation and Spectrum Assignment Algorithms in Elastic Optical Networks. *OFC/NFOEC Technical Digest*, v. 5, n. 10, p. A100 – A106, 2013. ISSN 1943-0620. Disponível em: <<https://www.osapublishing.org/jocn/abstract.cfm?uri=jocn-5-10-A100%5Cnhttp://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=6645104>>. Citado na página 36.

YU, Y.; ZHANG, J.; ZHAO, Y.; LI, H.; JI, Y.; GU, W. Exact performance analytical model for spectrum allocation in flexible grid optical networks. *Optical Fiber Technology*, v. 20, n. 2, p. 75–83, 2014. ISSN 10685200. Citado na página 20.

ZHANG, M.; SHI, W.; GONG, L.; LU, W.; ZHU, Z. Bandwidth defragmentation in dynamic elastic optical networks with minimum traffic disruptions. *IEEE International Conference on Communications*, p. 3894–3898, 2013. ISSN 15503607. Citado na página 29.

ZUKERMAN, M.; TAN, L.; WANG, H.; OUVEYSI, I. Efficiency-fairness tradeoff in telecommunications networks. *IEEE Communications Letters*, v. 9, n. 7, p. 643–645, 2005. ISSN 10897798. Citado na página 18.